THE UNIVERSITY OF ALBERTA

LASER SPECKLE INTERFEROMETRY: A STOCHASTIC INVESTIGATION

by Wolfgang Engler

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR
THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF PHYSICS

EDMONTON, ALBERTA
SPRING 2002

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-69704-5

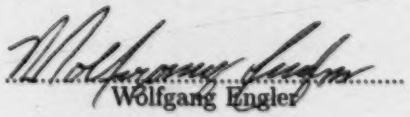
Canada

University of Alberta
Library Release Form

NAME OF AUTHOR:	Wolfgang Engler
TITLE OF THESIS:	Laser Speckle Interferometry: A Stochastic Investigation
DEGREE:	Master of Science
YEAR THIS DEGREE GRANTED:	2002

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



Wolfgang Engler

Address:
#1903-10135 Saskatchewan Drive
Edmonton, Alberta, Canada
T6E 4Y9

31 January 2002

The University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled *Laser Speckle Interferometry: A Stochastic Investigation* submitted by *Wolfgang Guenther Engler* in partial fulfilment of the requirements for the degree of *MASTER OF SCIENCE*.

.....
Dr. B. Schmitt (Supervisor)

.....
Dr. M. Heimpel (Physics)

.....
Dr. F. Marsiglio (Physics)

.....
Dr. M. Kouritzin (Mathematics)

January 18, 2002
.....
Date

Regulus laudandus est in
conservando iure iurando
—Marcus Tullius Cicero

for Daisy,
Lucky, Tina, Squeaky,
Daffy, Tuffy,
Frisky and Smoky

Abstract

Electronic speckle phase interferometry *(ESPI)* may be utilized in analysis of surface deformations. This thesis examines some benefits and drawbacks of interferometric deformation observation. A systematic approach demonstrates utility of stochastic methods in *ESPI* analysis. Statistics establish expected intensity distributions and interference effects for laser scatter. Purely statistical sampling errors are shown to dominate interferometric observation. Maximum likelihood analysis shows that constructive interference maxima, identified by global threshold, hold all information content; other interferogram measurements contribute negligible data weight in fitting procedures. An attempt to automate general deformation analysis is presented. The approach utilizes *singular value decomposition* to address interferometrically invisible displacement fields. Theoretical extensions allow for elimination of systematic errors from interferometer geometry measurement. Limitations of parametric deformation determination are explored through examination of a simple cantilever experiment. Physical single parameter models are shown to succeed where multi-parameter fits fail. Experimental results show consistency with theoretical *ESPI* precision limits.

Acknowledgements

While this thesis does record openly my facinations and efforts in pursuit of science, the foundation of the work rests deeply in a most human root. Only the support of good friends and colleagues has carried me, to keep my promise, by concluding this master's degree. To them I owe much gratitude.

Foremost, I would like to thank my supervisor Dr. D. Schmitt for his help and guidance in introducing me to the new and diverse field which is the subject of this thesis. I would also like to thank the rest of my committee, Dr. M. Heimpel, Dr. F. Marsiglio and Dr. M. Kouritzin, without whom these studies could not have been brought to completion.

Warmest thanks are due to Dr. Frank Weichman, without whose encouragement I might never have started this research. His continuing interest in my progress has been much appreciated. Many thanks go likewise to my friends and colleagues at *VisionSmart*, Daniel Kenway, Ann Marie Garcia, Ian Crowe, Cedar Mah and Terrance Van Gemert. Their patience and support have been of greatest help in finishing this work.

Finally, I would like to thank my good friends Alan Wilman, Malcolm Jones, and Daniel and Melanie Bilan for all the emotional support and balance they gave me throughout the course of these efforts.

Contents

I Theoretical foundations 1

1 Introduction 2

1.1 Objective 2

1.2 Background 6

2 Statistics and filters 13

2.1 Composite pixel intensity samples 13

2.1.1 Statistics for partially polarized light 14

2.1.2 Stochastic auto-correlations 17

2.1.3 Cross-correlation errors 20

2.1.4 Degrees of freedom 22

2.1.5 Cross-correlation weights 25

2.1.6 Systematic errors 27

2.2 Image filters 30

2.3 Data refits on raw and filtered images 32

3 Displacement fields **35**

3.1 Basic interferometry 35

3.2 General displacement fields 40

3.3 Stiff displacement fields 41

3.3.1 Rigid displacement fields 42

3.3.2 Polynomial displacement fields 46

3.4 Singular value decomposition 51

3.5 Source geometry adjustment 54

3.5.1 Geometry variation between images 55

3.5.2 Geometry variation from specifications 58

II Experiments **61**

4 Image processing **62**

4.1 Intensities 62

4.2 Filters 66

4.2.1 Video blur effect 67

4.2.2 Cross-correlation sampling errors 68

4.2.3 Location of maxima 71

5 Deformations **73**

5.1 Model data 73

5.1.1 Rigid translations 74

5.1.2 Rigid rotations 75

5.1.3 Rigid simulations 76

5.1.4 Polynomial deformations 80

5.1.5 Source geometry algorithms 83

5.2 Experimental data analysis 86

5.3 Cantilever deformation 96

5.3.1 Raw (5 × 5 sample) cross-correlations 110

5.3.2 Filtered cross-correlations 111

5.3.3 Theoretical deformations 112

5.3.4 Raw interferograms (scaled & offset) 113

5.3.5 Filtered interferograms (scaled & offset) 114

III Conclusions **115**

6 What has been shown? **116**

Bibliography 120

Appendices 124

A Probabilities 125

A.1 Distributions and densities 125

A.2 Means and medians 130

A.3 Convolutions 134

A.4 Central limit theorem 138

A.5 Simple stochastic process sums 141

A.6 Stochastic EM waves 146

A.6.1 Scalar wave approximation 146

A.6.2 Electromagnetic plane waves 151

A.6.3 Laser speckle intensity distribution 154

A.6.4 Laser speckle autocorrelation 157

A.6.5 Laser speckle phase interferometry 161

B Likelihoods 166

B.1 The Neyman-Pearson lemma 166

B.2 Maximum likelihood methods 172

B.3 Pearson's cross-correlation 174

B.4 General linear regression 177

C Computational methods **183**

C.1 Least squares stability analysis 183

C.2 The cantilever 186

D Notation **192**

List of Tables

3.1 Polynomial coefficients, fit rank and kernel size 49

4.1 Experimental distribution for $Q = \frac{N}{2} \frac{\sigma_{\rho\rho}^2}{(1-\rho^2)^2}$ (various samples) 69

5.1 Cantilever *SVD* fit results – extrapolated grid 82

5.2 Geometry tuning (mm scale) by rigid translation interferograms 85

5.3 Cantilever extremum filter 90

5.4 SVD (order=1) offset fit statistics: fit from all data 98

5.5 SVD (order=1) offset fit statistics: fit from end data 99

5.6 Constant vector offset fit statistics: fit from end data 100

5.7 Cantilever fit statistics: fit from all data 101

5.8 Cantilever fit statistics: fit from end data 102

List of Figures

2.1 Pixel cross-correlation effect: $n = 5$ & $N = 100$ 24

2.2 Experimental cantilever interferogram 29

3.1 Speckle interferometer layout 37

3.2 Effects of point displacement 37

4.1 Speckle pattern: single scattering source, $\sim 45^o$ incidence . . . 63

4.2 Polarized intensity vs. eqn 2.2 with $\mu = 15.8$ & $\sigma = 12.0$ . . . 64

4.3 Nonpolarized intensity vs. eqn 2.2 with $\mu = 30.0$ & $\sigma = 21.6$. 64

4.4 Blurred intensity vs. eqn 2.1 with $\mu = 108.9$ & $\sigma = 36.4$ 65

4.5 Experiment: single source vs. eqn 2.1 with $\mu = 14.23$ & $\sigma = 5.63$ 65

4.6 Experiment: dual sources vs. eqn 2.1 with $\mu = 34.4$ & $\sigma = 13.0$ 65

4.7 Blur effect: video blur enhances cross-correlation contrast . . . 67

4.8 Experimental distribution for $Q = \frac{N}{2} \frac{\sigma_{\rho\rho}^2}{(1-\rho^2)^2}$ (85 pixel samples) 69

4.9 Cross-correlation threshold for local maxima (raw data) 72

4.10 Cross-correlation threshold for local maxima (median data) . . 72

5.1 Translation: $\delta x = 2\mu m$ 74

5.2 Translation: $\delta y = 2\mu m$ 74

5.3 Translation: $\delta z = 2\mu m$ 74

5.4 Rotation: $\theta_x = 20n$radians 75

5.5 Rotation: $\theta_y = 20n$radians 75

5.6 Rotation: $\theta_z = 20n$radians 75

5.7 Model data interferometer parameters in *IFM* file format . . . 77

5.8 Cantilever interferogram (40g) – extrapolated theoretical . . . 80

5.9 Cantilever interferometer parameters in *IFM* file format . . . 81

5.10 Camera calibration (raw) 87

5.11 Camera calibration (perspective corrected) 87

5.12 View orientation (raw image) 87

5.13 View orientation (perspective corrected) 87

5.14 Cantilever reference image (0g) 89

5.15 Cantilever video image (15g) 89

5.16 Cross-correlation (15g) with (5 × 5) samples 89

5.17 Cross-correlation (perspective corrected) 90

5.18 After 7 × 7 median and radius 5.0 Gaussian 90

5.19 Identification of local extrema 90

5.20 Gradient by grey shade on local maxima (full image) 92

5.21 Gradient by grey shade on local maxima (right half image) . . 92

5.22 Theoretical interferogram: primary (maxima only) cantilever fit 92

5.23 Theoretical interferogram: primary (maxima only) *SVD* fit . . 92

5.24 Raw result with biquadratic scale and offset to cantilever fit . 93

5.25 Filtered result with biquadratic scale and offset to cantilever fit 93

5.26 Cross-section correction: raw data give true cosine values . . . 94

5.27 Theoretical secondary cantilever fit (raw phase unwound) . . . 95

5.28 Theoretical secondary cantilever fit (filtered phase unwound) . 95

5.29 Fit interferogram comparison: *SVD* vs. Cantilever 104

5.30 Theoretical peak location variance as function of threshold . . 107

5.31 Theoretical peak distribution with optimal thresholding 107

5.32 Cantilever deformation under 10g 110

5.33 Cantilever deformation under 20g 110

5.34 Cantilever deformation under 30g 110

5.35 Cantilever deformation under 40g 110

5.36 Cantilever deformation under 10g 111

5.37 Cantilever deformation under 20g 111

5.38 Cantilever deformation under 30g 111

5.39 Cantilever deformation under 40g 111

5.40 Cantilever deformation under 10g 112

5.41 Cantilever deformation under 20g 112

5.42 Cantilever deformation under 30g 112

5.43 Cantilever deformation under 40g 112

5.44 Cantilever deformation under 10g 113

5.45 Cantilever deformation under 20g 113

5.46 Cantilever deformation under 30g 113

5.47 Cantilever deformation under 40g 113

5.48 Cantilever deformation under 10g 114

5.49 Cantilever deformation under 20g 114

5.50 Cantilever deformation under 30g 114

5.51 Cantilever deformation under 40g 114

C.1 Cantilever model parameters in *CNT* file format 190

Part I

Theoretical foundations

Chapter 1

Introduction

1.1 Objective

The term *interferometry* encompasses a wide diversity of research techniques. These methods are all applications of the elementary wave phenomenon of interference. Interference of electromagnetic waves (light) was first explored by Thomas Young in 1801.[HECH 74] Modern applications utilize more exotic techniques, but produce simple fringe patterns which Young himself would recognize as wave effects. All such methods have the identical goal of obtaining fringe patterns representing interference effects between light from a pair of coherent light sources projected on a surface.

The study of interferometry, as examined in this thesis, derives as a specific application of the general interference principle. Microscopic deformations of a projection surface may be observed through interference effects between an illuminated view of the surface after deformation and recon-

struction of a reference view before deformation. As will become apparent in the following, this amounts to a rather elegant application of Young's experiment.

In applications of this method one obtains fringe patterns which represent interference effects of light from a coherent source pair. Observed fringe patterns derive from three-dimensional deformations of the projection surface between the observations. A key question arises in such examination. How do observed fringe patterns relate to specific deformation fields? The answer to this question should enable measurement of deformation vector fields on the surface.

In this thesis, the relationship between observed fringe interferograms and underlying deformation fields is subject to detailed study. Deformation measurement utility and inherent limitations to precision are explored, through application of the modern method of electronic speckle phase interferometry. Since the goal is to explore the limitations of real experimental measurement, all analysis techniques are programmed and explored in the versatile *C* language.[KERN 88]

Electronic speckle phase interferometry *(ESPI)* has been the state of the art in interferometry research in the last few decades.[DAIN 89][JONE 83] Several methods of interferogram production did in fact exist prior to the development of electronic image processing. Historical methods showed elegance and extravagance in design and function.[HECH 74][SCHM 89] Modern *ESPI* lacks in nostalgic grace, but also endows practical functionality on

this new measurement tool. In the present thesis, interferograms are observed by a modern method, involving digital analysis.

Traditional interferometry assumes the use of polarized light sources. Reference works on *ESPI* methods preserve this tradition.[DAIN 89][JONE 83] As introduction to a practical holographic interferometry, this thesis examines interferometry with nonpolarized scattering sources. The relationship between phase angle and shifting speckle pattern will be subjected to detailed study. Investigation of related intensity phenomena follows in a modern analysis of *stochastic* processes.

The method of *speckle interferometry* can be applied to measure micrometer scale (and perhaps smaller) displacements over a surface. The current thesis restricts investigation and measurement to a very simple interferometer design. Some advantages and detriments of *ESPI* and general dual-source interferometry are explored from such a context. Initially, one sets out to develop an algorithm solving the problem of deformation measurement. Along that path, one encounters a rich diversity of fundamental principles and restrictions defining speckle interferometry.

The key goal of this study remains to find interpretation of interferograms as representations of displacement fields. In solution of interferogram to deformation field inversion, an incomplete method is presented. Investigation will confirm that the problem is not invertible in a pure sense. No relationship exists for assigning a unique deformation field to an observed interferogram. One does however find a solution subject to some quasi-rigid constraints.

One finds, moreover, that theoretically inherent data noise restricts solution to relatively simple displacement fields.

The classical interference problem includes analysis of observed intensity fringes for unwinding (inversion) to continuous local phase angle. As the inverse cosine function is multi-valued, no unique assignment exists. Much research exists[JONE 83] into the problem of stable *phase-unwinding* algorithms. Most methods involve complex multi-interferogram measurements. Unwinding methods are unwieldy: they introduce series of new and often purely computational problems. *ESPI* interferograms derive from physical surface deformations. A method is presented, for simple data from single interferograms, which unwinds phase angle by extrapolation from a real deformation model.

Aside from simplicity, real model phase unwinding has the advantage of physicality: failure occurs precisely under conditions where quasi-rigid displacements cannot be measured, with such insensitivity deriving entirely from interferometer geometry. The presented method thus shows utility for identification of regionalized data noise, and as well, for finding real discontinuities such as fractures and fissures.

The relationship between quasi-rigid constraints and solution spaces yields new insight into the inversion problem. Limitations of the simple single observation interferometer are explored in this context. Extensions to basic deformation field measurement are investigated as well. An algorithm is presented, which affords a simple technique for evaluation of measurement

errors in scattering source geometry. Location correction may in fact be accomplished from observation of a single simple surface displacement.

The influence of perceived errors on interferometric results is examined in detail. Speckle interferometry involves observation of inherently noisy images. Methods for filtering visible noise from interferograms are demonstrated. Observation noise, however, stems from theoretical properties rather than from experimental errors. This reality imposes some rather severe restrictions regarding the information content of interferograms.

Theoretical and experimental research investigate the nature of inherent noise in speckle interferometry. Propagation of data noise, through analysis of displacement fields, finally imposes limitations on accuracy of deformation measurment. General results are presented, establishing the limitations of interferometric deformation measurement.

1.2 Background

The established field of interferometry has its foundation in the wave theory of light.[HECH 74][JACK 62] The phenomenon of source pair interference was introduced by Thomas Young at the dawn of the nineteenth century. Young's experiment resolved that light was in fact a wave phenomenon. Interference phenomena have been of scientific interest ever since.

Surface measurements involving wave-front reconstruction (holography) began more than half a century ago. Application of interference to hologra-

phy was demonstrated by *Gabor* in 1948, before the advent of lasers.[JONE 83] Practical holographic interferometry became a reality with the invention of coherent monochromatic light sources (lasers) and the subsequent 1962 study of *Leith* and *Upatnieks*.[DAIN 89]

Early work in holographic interferometry was dominated by the development of photographic techniques for reconstruction of multiple source holograms from multiply exposed film.[HECH 74] Old techniques involved careful acquisition and reconstruction of laser wave-fronts from rigidly mounted photographic plates. In practice, extremely clean optics were required to generate smooth wave-fronts: many interferometry techniques devoted great care toward *avoiding* laser speckle. Such methods were at once elegant but also impractical.

The speckle effect was first utilized by *Leendertz* in 1970.[LEEN 70] *Leendertz* demonstrated the utility of the Pearson's cross-correlation as a measure of phase shift. His original experiment involved manual alignment of a pair of photographic plates; effectively, this is equivalent to image cross-correlation. Today, the method may be more simply implemented via electronic image processing.

Electronic methods also allow the implementation of other powerful techniques. Some methods, for example, proposed by *Jones* and *Wykes* (1983) consist of pixel-by-pixel square or rectification of difference.[JONE 83] Such methods, amounting to the most basic forms of cross-correlation, are simple but highly sensitive to noise. Image fringes appear more as textural fea-

tures than interference phase amplitudes in resulting interferograms. More complex algorithms may be devised for better results. For example, pixel averaging before applying a square root to rectified differences also generates fringe images. Systematic methods generally improve interferograms by respecting the nature of the laser speckle effect.

Amongst sophisticated interferometric algorithms, a single method stands out in elegant simplicity. Interferograms of excellent quality obtain from the method of image cross-correlation.[JONE 83] This approach, introduced by *Jones and Wykes*, has another distinct advantage: image noise level may be traded off against spatial resolution, by altering spatial dimensions of statistical sample regions. Investigations of this thesis restrict study of speckle interferometry to the robust Pearson's cross-correlation method.

In order to facilitate development of physical interferometry concepts, one requires the statistical foundations for this observational theory. Statistical concepts involved in cross-correlataion are well established. Work throughout this thesis utilizes conventional definitions and notations. The interested reader is referred to a detailed development in the chapter A & B appendices.

Given random variable ordered pair (x, y), and an empirical sample distribution of such pairs over independent trials, one may estimate variances σ_{xx}^2 and σ_{yy}^2 and shared covariance σ_{xy}^2; and one may write the theoretical Pearson's cross-correlation:[FREU 62]

$$\rho(x,y) \;=\; \frac{\sigma_{xy}^2}{\sqrt{\sigma_{xx}^2\,\sigma_{yy}^2}} \tag{1.1}$$

For real empirical samples, cross-correlation has maximum likelihood estimate expressed, per Appendices B.2 and B.3, as follows:

$$\rho(x,y) \;=\; \frac{N\sum xy \;-\; \sum x \sum y}{\sqrt{(N\sum xx \;-\; \sum x \sum x)\,(N\sum yy \;-\; \sum y \sum y)}} \tag{1.2}$$

From the non-negative semi-definite property of the correlation tensor, one finds for any bivariate distribution $P(x,y)$, that Pearson's cross-correlation observes boundedness:

$$-1 \;\leq\; \rho(x,y) \;\leq\; +1 \tag{1.3}$$

A few special cases are of note. If random variables x and y are independent, then $\rho(x,y) = 0$. At the other extreme, $\rho(x,y) = \pm 1$ corresponds to an exact linear relationship: $ax + by = c$. The relationship holds exactly, except possibly on a set of zero (probability) measure. For the special case where $\langle x \rangle = \langle y \rangle$ and $\sigma_{xx}^2 = \sigma_{yy}^2$, perfect correlation $\rho(x,y) = +1$ implies $x = y$ exactly on a set with fully unity for (probability) measure.

In application to laser speckle interferometry, justification for use of Pearson's cross-correlation derives from straight forward application of electromagnetism and interference. One may examine the phenomenon with various degrees of subtlety and rigour. The following gives an overview of simple documented results. Detailed development of the stochastic basis for the theory of speckle interferometry is laid out in Appendix A.6.

Interference between plane waves of coherent frequency light, with parallel propagation and single polarization, models as a scalar wave equivalent.

Superposition of a pair of simple coherent waves with intensities i_A & i_B and relative phase offset ϕ (in radians) results in following net intensity:

$$I(i_A, i_B; \phi) \;=\; i_A \,+\, i_B \,+\, 2\sqrt{i_A i_B}\cos\phi \qquad (1.4)$$

A simple experiment is devised as follows: suppose two known intensity distributions i_A & i_B are generated and sampled, with respective phase shifts α & β applied to each. If one replaces normalized sums by expectations in the Pearson's cross-correlation sample formula (equation 1.2), one recovers the global formula (equation 1.1) and obtains:

$$\rho(I(i_A, i_B; \phi+\alpha), I(i_A, i_B; \phi+\beta)) \;=\; 1 + \kappa(\{i_A\}, \{i_B\})[\cos(\alpha-\beta)-1] \quad (1.5)$$

One makes a few reasonable assumptions for statistical calculation: random variables i_A, i_B, and uniformly distributed angle ϕ must be independent. Under such conditions one finds, in terms of averages $\langle i_A \rangle$ & $\langle i_B \rangle$ and variances $\sigma^2_{i_B i_B}$ & $\sigma^2_{i_A i_A}$, a general result holds for such simple waves:[JONE 83]

$$\kappa(\{i_A\}, \{i_B\}) \;=\; \frac{2\langle i_A \rangle \langle i_B \rangle}{\sigma^2_{i_A i_A} + 2\langle i_A \rangle \langle i_B \rangle + \sigma^2_{i_B i_B}} \qquad (1.6)$$

One can in fact improve this expression. *Goodman* showed in 1975 (see Appendix A.6) that random distribution of point intensity, due to laser speckle from single polarized source, or scalar model equivalent, follows an *exponential* probability density:[GOOD 89]

$$P(i) \;=\; \frac{1}{\langle i \rangle}\,\exp(-\frac{i}{\langle i \rangle}); \quad \text{with} \quad \mu = \langle i \rangle \quad \& \quad \sigma^2_{ii} = \langle i \rangle^2 \qquad (1.7)$$

Goodman's distribution allows for simplification of formula 1.6, yielding in the simple model a particular result:

$$\kappa(\{i_A\},\{i_B\}) \;=\; \frac{2\langle i_A\rangle\langle i_B\rangle}{(\langle i_A\rangle+\langle i_B\rangle)^2} \tag{1.8}$$

One should note with caution that the foregoing results from equations 1.4 through 1.8 describe only coherent plane wave light of parallel propagation and with single polarization. This scalar model is illustrative but in no sense reasonable for any *ESPI* experimental setup. Equation 1.4 assumes a single polarization direction with colinear light sources and observation point. One can in principle arrange polarizing filters to create single polarization scattering conditions locally near some reference point; exotic filters would be required though, to arrange such conditions globally over a reference frame. Colinearity of light sources and reference points cannot, however, be sensibly attained or even approximated over any extended reference surface. One concludes that the simple model is wrong for any reasonable interferometer design.

Surprisingly, the foregoing simple results may still be extrapolated to non-polarized light sources. A demonstration follows in the next chapter, giving results with similar properties, but showing angular variation across the reference surface. Angular variability complicates the effect, but does not affect its utility as a measurement tool. In practice, one requires only a method which allows observation of fringe patterns for subsequent inversion into displacement fields. One finds in the foregoing, for point intensity observation

of polarized light, that if statistical moments of i_A & i_B vary slowly on a region of interest, one may simplify formula 1.5 in terms of slowly varying scalar fields:[JONE 83]

$$\rho(\, I(\phi+\alpha), I(\phi+\beta)\,) \;=\; f(x,y) \;+\; g(x,y)\cos(\alpha-\beta) \qquad (1.9)$$

$$\text{where} \quad |\nabla f| \ll |\nabla(\alpha-\beta)| \quad \text{and} \quad |\nabla g| \ll |\nabla(\alpha-\beta)|$$

The empirical interpretations expressed in equations 1.5 and 1.9 show surprising applicability: interferograms may be observed over wide variations in source location and composition. Images show resilience, as well, under adjustment of other experimental minutae. For example, experimental observations show insensitivity to variation of speckle size and optical blur. Theoretical properties of cross-correlation predicate against adverse effects from such experimental observables, and one in fact often observes improvement of experimental interferometric image quality when aberrations are artificially introduced.

Chapter 2

Statistics and filters

2.1 Composite pixel intensity samples

The laser speckle phenomenon studied and employed in this work was generated by dual source scatter, with no attempt to repolarize either sources or observed light. Interferogram observation showed great resilience under other significant experimental deviations from simple theoretical models. Many elements of experimental setup in practice required surprisingly little attention. Quality of speckle patterns was seen to be largely unaffected by camera location, precise source location or source composition. Even camera focus and background light were found to have little effect on fringe observation. Aside from apparatus rigidity, the principal ingredients for generation of useful correlation interferograms were found to be all in the laser: one simply requires sufficient intensity and coherence from a laser source.

The resilience of the speckle cross-correlation phenomenon under patho-

logical conditions is in fact predicted by theory. The reasons behind the resilience to changing from polarized to nonpolarized sources and the tolerance to loss of focus are physically unrelated. Mathematically, however, they stem from the same root: intensity cross-correlation shows invariance under a class of statistically similar effects.

2.1.1 Statistics for partially polarized light

The polarized point intensity due to *Goodman* (equation 1.7) is the lowest order ($\alpha = 1$) term of the more general gamma density. Conventionally parametrized by real $\alpha > -1$ and $\beta > 0$, the gamma density may be expressed over random variable x as follows:[FREU 62]

$$\gamma(x;\alpha,\beta) = \frac{1}{\beta^{\alpha}\,\Gamma(\alpha)}\,x^{\alpha-1}\exp(-\frac{x}{\beta}) \tag{2.1}$$

$$\text{with mean} \quad \mu = \alpha\beta \quad \text{and variance} \quad \sigma_{xx}^2 = \alpha\beta^2$$

Physically, the experimental setup is concerned with observation of light intensity over an angular field of view, subdivided into pixels of finite size. Net observed light intensity may be viewed as a sum of contributions. The contributions are due to polarizations at a point, and due to independent point intensities which may contribute to the same pixel. Since the probability of a sum amounts to a simple convolution, one expects that theoretical light intensity should follow properties of gamma distributions under convolution. Derivation of relevant convolutions can be found in Appendix A.3.

One may most immediately utilize gamma convolution formulas, to investigate theoretical intensity distribution for nonpolarized light from its polarization mean intensities. One may assume polarizations $\top$ and $\perp$ correspond to independent random variables, and thus infer principal axes where $\sigma^2_{i_\perp i_\top} = 0$. If such independent intensity components are distributed as exponential densities (equation 1.7) viewed as lowest order ($\alpha = 1$) gamma densities (equation 2.1), then one may utilize gamma convolution (per equations A.36 and A.37) to obtain a physical result:

$$P(i) = \begin{cases} \frac{1}{1-\frac{\langle i_\top \rangle}{\langle i_\perp \rangle}} \gamma(i; 1, \langle i_\perp \rangle) + \frac{1}{1-\frac{\langle i_\perp \rangle}{\langle i_\top \rangle}} \gamma(i; 1, \langle i_\top \rangle) & \langle i_\perp \rangle \neq \langle i_\top \rangle \\ \gamma(i; 2, \frac{1}{2}(\langle i_\perp \rangle + \langle i_\top \rangle)) & \langle i_\perp \rangle = \langle i_\top \rangle \end{cases} \tag{2.2}$$

$$\begin{aligned} \langle i \rangle &= \langle i_\perp \rangle + \langle i_\top \rangle \quad \text{and} \\ \sigma^2_{ii} &= \langle i_\perp \rangle^2 + \langle i_\top \rangle^2 \quad \text{or} \\ \langle i_\perp \rangle &= \frac{1}{2}\left(\langle i \rangle + \sqrt{2\sigma^2_{ii} - \langle i \rangle^2} \right) \quad \text{and} \\ \langle i_\top \rangle &= \frac{1}{2}\left(\langle i \rangle - \sqrt{2\sigma^2_{ii} - \langle i \rangle^2} \right) \end{aligned}$$

The distribution $P(i)$ describes light intensity from a single, possibly partially polarized source. Note that purely nonpolarized light has variance σ^2_{ii} only half as large as purely polarized light of the same intensity. (Compare *Goodman's* result, equation 1.7 with the above expression in equation 2.2).

The alert reader will note that while the preceding derivation outline is quite devoid of physical content, the result does predicate symmetric intensity

tensor $\mathcal{I}$, which may be expressed in 3 dimensions as follows:

$$(\hat{k}\cdot\hat{n})\mathcal{I} \;=\; \begin{pmatrix} i_{xx} & i_{xy} & i_{xz} \\ i_{yx} & i_{yy} & i_{yz} \\ i_{zx} & i_{zy} & i_{zz} \end{pmatrix} \;=\; R \begin{pmatrix} i_{\perp} & 0 & 0 \\ 0 & i_{\top} & 0 \\ 0 & 0 & 0 \end{pmatrix} \tilde{R} \qquad (2.3)$$

Surface incidence derives from (ray) axis of propagation unit vector $\hat{k}$ and surface unit normal vector $\hat{n}$ as a simple inner product. Rotation R gives the conversion from principal axes coordinates. The null eigen-mode corresponds to the (ray) axis of plane wave propagation. One may immediately deduce the following coordinate-independent results:

$$\langle i \rangle \;=\; trace(\mathcal{I})\,(\hat{k}\cdot\hat{n}) \quad \& \quad \sigma_{ii}^2 \;=\; trace(\mathcal{I}\mathcal{I})\,(\hat{k}\cdot\hat{n})^2 \qquad (2.4)$$

The cross-correlation coefficient for partially polarized light may be evaluated in terms of tensors $\mathcal{I}_A$ & $\mathcal{I}_B$, due to the respective sources A & B, with regard to propagation axes along (unit) rays $\vec{k}_A$ & $\vec{k}_B$, and with reference to the local (average) surface unit normal $\hat{n}$. The result of detailed derivation, outlined in Appendix A.6, gives $\rho(I_A, I_B)$ following the form of equation 1.5. The general partially polarized result appears as follows:

$$\kappa(\{i_A\},\{i_B\};\hat{n}) \;=\; (1 \;+\; \frac{\mathcal{T}(\{i_A\},\{i_B\};\hat{n})}{\mathcal{S}(\{i_A\},\{i_B\};\hat{n})})^{-1}; \quad \text{where} \qquad (2.5)$$

$$\mathcal{S}(\{i_A\},\{i_B\};\hat{n}) \;=$$

$$(\tilde{n}\mathcal{I}_A\vec{n})(\tilde{k}_A\mathcal{I}_B\vec{k}_A) \;+\; 2\cdot(\tilde{n}\mathcal{I}_A\vec{k}_B)(\tilde{k}_A\mathcal{I}_B\vec{n}) + (\tilde{k}_B\mathcal{I}_A\vec{k}_B)(\tilde{n}\mathcal{I}_B\vec{n}) \;+$$

$$[\tilde{n}\cdot(\vec{k}_A+\vec{k}_B)]^2 trace(\mathcal{I}_A\mathcal{I}_B) - [\tilde{n}\cdot(\vec{k}_A+\vec{k}_B)][\tilde{n}(\mathcal{I}_A\mathcal{I}_B+\mathcal{I}_B\mathcal{I}_A)(\vec{k}_A+\vec{k}_B)]$$

$$\mathcal{T}(\{i_A\},\{i_B\};\hat{n}) =$$
$$\frac{1}{2}(\tilde{n}\cdot\vec{k}_A)^2\cdot[trace^2(\mathcal{I}_A)+3trace(\mathcal{I}_A^2)] +$$
$$\frac{1}{2}(\tilde{n}\cdot\vec{k}_B)^2\cdot[trace^2(\mathcal{I}_B)+3trace(\mathcal{I}_B^2)]$$

One may now derive the simple result as a limiting case. The polarized limit of equation 2.5, with parallel polarizations from sources A and B, and both propagation axes parallel to the surface normal yields exactly the requisite reduction.

$$\kappa(\{i_A\},\{i_B\};\hat{n}) = (1 + \frac{\mathcal{T}(\{i_A\},\{i_B\};\hat{n})}{\mathcal{S}(\{i_A\},\{i_B\};\hat{n})})^{-1} = \frac{2\cdot\langle i_A\rangle\langle i_B\rangle}{(\langle i_A\rangle+\langle i_B\rangle)^2} \quad (2.6)$$

One notes again that the simplified $\kappa(\{i_A\},\{i_B\};\hat{n})$ of equation 1.8 results only when experimentally unrealistic conditions are precisely met. Functions $\mathcal{S}(\{i_A\},\{i_B\};\hat{n})$ and $\mathcal{T}(\{i_A\},\{i_B\};\hat{n})$ do in fact generally appear quite complex. Since tensors $\mathcal{I}_A$ and $\mathcal{I}_B$ are expected to vary slowly on the surface, and since $\vec{k}_A$, $\vec{k}_B$ and $\vec{n}$ are analytic and slowly varying, equation 2.5 does however remain of the requisite type, as expressed in equations 1.5 and 1.9.

2.1.2 Stochastic auto-correlations

Statistical treatment of light intensity, as presented in the preceding section, relies on the presumption of interference between simple plane waves. The phenomenon of laser speckle does not seem, apriori, to satisfy this presumption: a noisy source of electromagnetic radiation expanded in vector spherical harmonics[JACK 62] yields a noisy speckle pattern only if high order terms

are of significant magnitude. The question of validity of the local plane wave approximation can best be put as follows. Does a limit exist, regulating perceived speckle size and setting a scale at which the plane wave approximation becomes valid?

The assumption of a stochastic source resolves this question, and in fact allows for assessment of relevant autocorrelations. One immediately notes that a stochastic source sensibly approximates a real source. One expects that scattering at the surface stems from random distribution of many micro-sources. Micro-sources should show random spatial distribution over the physical scattering target, and include also random phase offsets. In the limit as sheer quantity of micro-sources becomes unbounded, net observables attain stochastic characteristics. One may therefore model the resultant electromagnetic field by means of stochastic processes and Wiener integration.

One may utilize stochastic methods to evaluate Pearson's cross-correlation for spatially separated intensity measurements. General examination of requisite stochastic processes is presented in Appendices A.5 and A.6, with the particular derivation in Appendix A.6.4. For the case of a single Gaussian stochastic source, with observed source intensity proportional to $G_3(\vec{x}; \vec{\mu}_{\text{sce}}, \vec{\vec{\sigma}}^2_{\text{sce}})$, (per equation B.17,) one may write the result as follows:

$$\rho(i(\vec{x}), i(\vec{y})) \sim \exp\left(-\frac{1}{2} (\vec{x}-\vec{y}) \cdot \sigma^{2\ -1}_{\text{speck}} \left(\frac{1}{2}(\vec{x}+\vec{y}) - \vec{\mu}_{\text{sce}}; \sigma^2_{\text{sce}}, \lambda\right) \cdot (\vec{x}-\vec{y}) \right) \quad (2.7)$$

$$\sigma^{2\ -1}_{\text{speck}}(\vec{r}; \sigma^2_{\text{sce}}, \lambda) = 2 \left(\frac{2\pi}{\lambda |\vec{r}|}\right)^2 (1 - \hat{\vec{r}} \otimes \hat{\vec{r}})\, \sigma^2_{\text{sce}}\, (1 - \hat{\vec{r}} \otimes \hat{\vec{r}}) \quad \text{where}$$

$$\hat{\vec{r}} = \frac{\vec{r}}{|\vec{r}|} \quad \text{asymptotically for} \quad \lambda \ll \sqrt{trace(\sigma^2_{\text{sce}})} \ll r$$

Equation 2.7 holds regardless of polarization states of incoming and outgoing light at the scattering source. Since most conventional laser sources moreover generate beams of rated Gaussian (cross-section) intensity, one expects the result to hold validity under realistic experimental conditions.

From the stochastic result, or the related autocorrelation function, one may infer that scattering source size regulates speckle dimension. Micro-source characteristics would have to be quite pathological indeed, to generate coherent multipole radiation with so great an angular variability. Since one cannot sensibly expect such pathological micro-sources, and since experimental speckle dimension accords well with the stochastic description, (per equation 2.7,) one concludes that the stochastic asymptotic plane wave model holds validity for the laser speckle phenomenon.

One may draw a predictive inference regarding scattering source properties and speckle nature. The stochastic model makes no explicit use of surface qualities, such as composition and roughness. Provided surface qualities suffice to approximate stochastic scatter, theory predicts that other source qualities should have little effect on the nature of speckle.

A statistical inference follows as well from the stochastic treatment. Pixel size and speckle dimension are often aligned[ENNO 89][JONE 83] in the literature. Theoretically, no requirement exists for a strict ordering relationship. One requires only that statistical intensity fluctuations be observable. Rigorous examination (see Appendix A.6.5) shows that pixel scale does alter perceived cross-correlations, by respecting equation 1.5, with modification

only to scaling function $\kappa(\{i_A\},\{i_B\};\hat{n})$. Intuitively however, observed pixel cross-correlations should approximate underlying point intensity equivalents.

Consider joint trial variable pair x_n & y_n, over independent trials by index n, with respective means $\bar{x}_n$ & $\bar{y}_n$. For fixed weights $w_n \in \mathbf{R}$, one may express composites X & Y with following characteristics:

$$\begin{aligned} X = \sum_n w_n x_n \quad \& \quad Y = \sum_n w_n y_n \quad \longrightarrow \\ \bar{X} = \sum w_n \bar{x}_n \quad \& \quad \bar{Y} = \sum w_n \bar{y}_n \\ \begin{bmatrix} \sigma^2_{XX} & \sigma^2_{XY} \\ \sigma^2_{YX} & \sigma^2_{YY} \end{bmatrix} = \sum_n w_n^2 \begin{bmatrix} \sigma^2_{x_n x_n} & \sigma^2_{x_n y_n} \\ \sigma^2_{y_n x_n} & \sigma^2_{y_n y_n} \end{bmatrix} \end{aligned} \tag{2.8}$$

If all measurement covariances are scalar multiples of an identical tensor, $\vec{\sigma}^2_n = s_n \vec{\sigma}^2$, then cross-correlations for contributions and for composites must equate as $\rho(x_n, y_n) = \rho(X, Y)$.

One may regard composite pixels in a similar context, as combinations of independent point intensities. Pixel intensities should thus show cross-correlations equivalent to those shared by all local asymptotic plane-wave pair contributions. While magnitudes of intensity means and variances may fluctuate, any suitable optical arrangement should extract similar cross-correlation values.

2.1.3 Cross-correlation errors

Pearson's cross-correlation as expressed in formula 1.2 is based on a statistical sample. Being a measurement, this quantity is subject to standard error. Evaluation of the linearized error term (see Appendix B.3) gives for the

cross-correlation variance the following result:

$$\sigma^2_{\rho\rho} \;=\; \frac{2}{N}\,(1-\rho^2)^2 \tag{2.9}$$

Numerical constant N refers to real degrees of freedom inside the sample, rather than pixel number. These quantities should however ordinarily be similar. Surprisingly, the variance may be expressed solely in terms of the quantity itself, without reference to moments of underlying measurements.

Quite sensibly, however, under equation 2.9, cases of perfectly correlated (error free) data with $\rho = \pm 1$ give error $\sigma^2_{\rho\rho} = 0$, and uncorrelated data with $\rho = 0$ give maximal error $\sigma^2_{\rho\rho} = \frac{2}{N}$. In terms of interferograms one infers, with reference to formula 1.9, that regions of constructive interference should have high and quiet signals, whereas regions of destructive interference should give low and noisy measurements. With view to equation 1.5, such respective conditions correspond to interference angles where $\cos(\alpha-\beta) = +1$ constructively and $\cos(\alpha-\beta) = -1$ destructively.

One may predict a measurable result in terms of equation 2.9, by considering the linearized error term for $\sigma^2_{\rho\rho}$. Direct evaluation shows:

$$\begin{aligned}
\frac{\delta\sigma^2_{\rho\rho}}{\sigma^2_{\rho\rho}} \;&\sim\; \tfrac{-4\,\rho\,\delta\rho}{1-\rho^2} \;\longrightarrow \\
var(\nu\,\frac{\sigma^{2\ \text{linear}}_{\rho\rho}}{\sigma^2_{\rho\rho}}) \;&\sim\; 16\nu\,\tfrac{\sigma^2_{\rho\rho}}{(1-\rho^2)^2}\,\rho^2 \sim\; 32\,\frac{\nu}{N}\,\rho^2 \quad \text{whereas} \\
var(\nu\,\frac{\sigma^{2\ \text{sample}}_{\rho\rho}}{\sigma^2_{\rho\rho}}) \;&\sim\; 2\nu \quad (\sim \chi\text{-dist}) \;\longrightarrow \\
\frac{var(\sigma^{2\ \text{linear}}_{\rho\rho})}{var(\sigma^{2\ \text{sample}}_{\rho\rho})} \;&\sim\; \tfrac{16}{N}\,\rho^2
\end{aligned} \tag{2.10}$$

By inference, so long as the ratio stays negligible, either as a limit (ie. $1 \gg \frac{16}{N}$), or less restrictively in average (ie. $1 \gg \frac{16}{N}\langle\rho^2\rangle \sim \frac{5}{N}$), one may define by random variable Q the sample ratio, such that νQ has approximate (chi-square) χ^2-distribution in ν degrees of freedom:

$$Q(\rho, \sigma^2_{\rho\rho}) \;=\; \frac{\sigma^{2\,\text{sample}}_{\rho\rho}}{\sigma^{2\,\text{linear}}_{\rho\rho}} \qquad (2.11)$$

where the χ^2-distribution in ν degrees of freedom derives as following special case of the more general γ-distribution:[FREU 62]

$$\chi(x;\nu) \;=\; \gamma(x;\, \alpha = \frac{\nu}{2},\, \beta = 2) \quad \longrightarrow \quad \mu = \nu \;\&\; \sigma^2 = 2\nu \qquad (2.12)$$

2.1.4 Degrees of freedom

An interesting consequence follows from equations 2.11 and 2.12: one may measure the pixel density of perceived degrees of freedom for a cross-correlation interferometer. One has only to interpret the experimental distribution of random variable Q in terms of the general gamma distribution.

In counting degrees of freedom, one identifies a key proviso. While each measurement does constitute a true increment in degrees of freedom, one has not yet accounted for cross-correlations amongst neighbouring measurements. One may express for pixels whose sample regions overlap, following approximate cross-correlation result:

$$\sigma^2_{a,b}(\rho) \;=\; \frac{N_{\text{shared}}(a,b)}{N_{\text{sample}}}\,\sigma^2_{\rho\rho} \qquad (2.13)$$

Cross-correlations amongst sample regions are in fact significant. Their

effect may however be approximated as a pure reduction in perceived degrees of freedom. Viewing σ^2 in terms of a basis of its eigenvectors, approximations may be compared to the true tensor form. If measurements were truly independent, sharing identical variance, their number would simply count:

$$\frac{\nu}{N} = \frac{trace^2(\sigma^2)}{N\ trace(\sigma^{2\,2})} \tag{2.14}$$

One obtains a pair of results for $n \times n$ cross-correlation samples along a line of length $N \gg n$ where $\sigma^2_{\rho\rho}$ is constant. If equation 2.14 holds validity for the experimental correlation tensor, a simple result follows:

$$\frac{\nu}{N} = \frac{n^2}{2(\sum_{k=1}^{k=n} k^2) - n^2} = \frac{3n}{2n^2+1} \tag{2.15}$$

Another lower bound in detectable degrees of freedom derives from an intuitive observation. One expects minimally, that nonoverlapping regions count significantly:

$$\frac{\nu}{N} = \frac{1}{n} \tag{2.16}$$

The apriori assumption of independence for each pixel as a degree of freedom clearly differs greatly from these lower approximations.

As example for the overlap effect, one may examine the case of 5×5 cross-correlation samples, with linear single pixel incrementation along a vertical line. Cross-correlation interferogram images generated throughout this thesis follow precisely this prescription. In regions where cross-correlation ρ is relatively constant, one expects measurements along the line to be represented

according to equation 2.13 as follows:

$$\frac{5}{\sigma^2_{\rho\rho}}\,\sigma^2 = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & \\ 4 & 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & \\ 3 & 4 & 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 & \\ 2 & 3 & 4 & 5 & 4 & 3 & 2 & 1 & 0 & 0 & \\ 1 & 2 & 3 & 4 & 5 & 4 & 3 & 2 & 1 & 0 & \cdots \\ 0 & 1 & 2 & 3 & 4 & 5 & 4 & 3 & 2 & 1 & \\ 0 & 0 & 1 & 2 & 3 & 4 & 5 & 4 & 3 & 2 & \\ 0 & 0 & 0 & 1 & 2 & 3 & 4 & 5 & 4 & 3 & \\ 0 & 0 & 0 & 0 & 1 & 2 & 3 & 4 & 5 & 4 & \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & 3 & 4 & 5 & \\ & & & & \vdots & & & & & & \ddots \end{bmatrix} \tag{2.17}$$

The real eigenvalue distribution, (in descending order,) for cross-correlation images satisfying equation 2.17, appears in figure 2.1. Trace (equation 2.15) and minimum dimension (equation 2.16) approximations have been superimposed. While the approximations remain poor, real distribution of sample

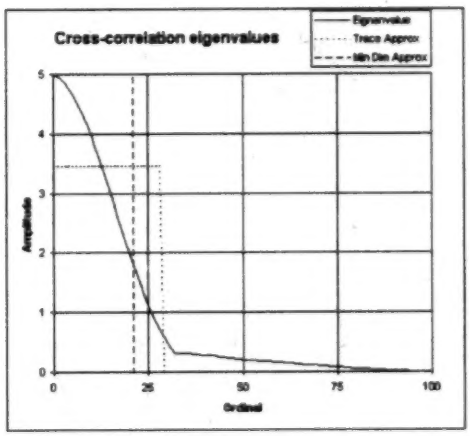


Figure 2.1: Pixel cross-correlation effect: $n = 5$ & $N = 100$

variance $\sigma^2_{\rho\rho}$ should match better with the approximations than with the initial presumption of total independence.

The presumption of minimal degrees of freedom seems to fit real eigenvalues better than does the trace-based approximation. When imposing a χ^2-distribution on a data sample, one expects, perceived degrees of freedom ν should approximate equation 2.16 better than 2.15, in modelling localized measurement correlation. While the prediction is weak, it nonetheless constitutes an observable.

2.1.5 Cross-correlation weights

In fits of surface deformations to interferograms, one should attempt to weight data so as to give an optimal fit. Because constructive and destructive interference show polar extremes in noise levels, low and high limits respectively, as demonstrated in subsection 2.1.3, one may in fact assign optimal weights via equation 2.9. Optimal measurement weighting is given, as outlined in equation A.16 by inverse variance. Consider equation 1.5 with function $\kappa((\{i_A\},\{i_B\})$ locally constant in the neighbourhood of a local maximum. Provided relative phase is locally smooth with almost constant gradient, one finds along that gradient:

$$\begin{aligned} \rho(x) &= 1+\kappa\,(\cos\phi-1) \;\sim\; 1-\frac{1}{2}\kappa\,\phi^2 \\ \sigma^2_{\rho\rho} &\sim \quad \kappa^2\,\sin^2\phi\;\sigma^2_{\phi\phi} \;\sim\; \kappa^2\,\phi^2\;\sigma^2_{\phi\phi} \end{aligned} \tag{2.18}$$

$$\sigma^2_{\rho\rho} = \frac{2}{N}(1-\rho^2)^2 \sim \frac{2}{N}\phi^4 \longrightarrow$$
$$\sigma^2_{\phi\phi} \sim \frac{2}{N\kappa^2}\phi^2$$

The optimal weight of a region amounts to the relative measure of that region. Simple integration thus yields regional (linearized) optimal weight. Near a local maximum, one finds the following asymptotic form for a closed linear interval:

$$W([a,b]) \sim \frac{1}{k}\int_{\phi(a)}^{\phi(b)} \frac{d\phi}{\sigma^2_{\phi\phi}} \sim \frac{N\kappa^2}{2k^2}\left(\frac{1}{\phi(a)} - \frac{1}{\phi(b)}\right) \qquad (2.19)$$
$$\text{where} \quad \phi \sim k\cdot(x-c) \quad \text{for } (x+c)\in[a,b]$$

One observes that weights near the local maximum dominate over any other region, beyond any proportionality constant. One concludes that observations from suboptimal regions fail to add new information content to fits involving perfectly correlated data points. Symbollically, one may express the result in set notation and write:

$$W(\phi^{-1}(0)) \gg W([a,b]-\phi^{-1}(0)) \longrightarrow \text{let } w(x) = \begin{cases} \frac{\kappa^2}{k^2} & |\rho| = 1 \\ 0 & |\rho| \neq 1 \end{cases} \qquad (2.20)$$

Optimal weighting results only when data are restricted to the kernel $\sigma^2_{\rho\rho} = 0$, or simply to where $\rho = \pm 1$. Weighting of local maxima follows the prescription of equation 2.20.

One might note that the variance in equation 2.18 fails to include a local term due to interferogram pixelation. An independent term may be added, expressing such a nonlocal effect, in the following form:

$$\sigma^2_{\phi\phi} \sim \sigma^{2\,0}_{\phi\phi} + \frac{2}{N\kappa^2}\phi^2 \qquad (2.21)$$

If, for example, one regards constant term $\sigma^{2\,0}_{\rho\rho}$ as caused by limitation of the local maximum to a box of n contiguous pixels, then uniform distribution (per equation A.11) gives $\sigma^{2\,0}_{\phi\phi} = (nk)^2/12$.

One may accordingly express weights associated with equation 2.21 in the form of an associated (Cauchy density) distribution:[FREU 62]

$$\begin{aligned} c(\phi;\Phi) &= \tfrac{\Phi}{\pi}\,(\Phi^2+\phi^2)^{-1}; \quad \Phi = \sqrt{\tfrac{1}{2}\,N\kappa^2\,\sigma^{2\,0}_{\phi\phi}} \\ C(\phi;\Phi) &= \textstyle\int_{-\infty}^{\phi} d\phi'\; c(\phi';\Phi) \quad = \dfrac{1}{2} + \dfrac{1}{\pi}\,\tan^{-1}\dfrac{\phi}{\Phi} \end{aligned} \tag{2.22}$$

The Cauchy density $c(\phi;\Phi)$ has conventional parameter Φ, and $C(\phi;\Phi)$ gives the associated integrated probability.

One finds 50% weight accumulation within $1 \times \Phi$ of the peak, and more than 70% within $2 \times \Phi$, and more than 84% within $4 \times \Phi$. Experimental values may be approximated as $N \sim 25$ (sample population), $\lambda \sim 50$ (interferogram pixel wavelength), $\kappa \sim 0.5$ (from equation 1.5), and $n \sim 2$ (pixel box localizing maximum). Such values give resulting $\Phi \sim 12.8\%$ of a radian, or more concretely, about 1.0 pixel. One concludes that, for regions away from local maxima, data weights remain negligible.

2.1.6 Systematic errors

Experimental cross-correlations are subject to systematic measurement errors. Real data, for example, should almost never yield a sample of identically constant phase offset. In particular, one expects greatest unlikelihood for choosing a perfect sample, with truly optimal cross-correlation $\rho = \pm 1$.

Errors due to sample size may however be mitigated by reduction of sample dimensions.

One must contend as well with video signal errors. Two primary sources of signal error were identified. First, true signal measurement noise from video signal, digitization or laser decoherence must be considered. Noise decorrelates from any signal, including itself under time delay. Its net effect is to scale cross-correlations by a uniform factor. Second, surface texture is illuminated by ambient and scattered laser light. Texture here includes all purely local lighting effects visible at the surface, including variabilities in colouration, roughness, reflectivity and light incidence. Texture should correlate almost perfectly. One notes as special case, that perfectly correlated (ie. $\rho = +1$) signals sum into perfectly correlated sums. Signal errors due to texture thus should alter all other measured cross-correlations slightly, but leave constructive interference maxima entirely unaffected.

One may express the cross-correlation of perceived intensity signals X and Y in terms of actual signals x, y and noise signal n and texture t, as well as two correlation coefficients shared by x and y in a region well centred between the scattering sources. Speckle intensity has local variance σ^2_{ss}, which applies for both signals x and y in the central region, and covariance σ^2_{st} with texture, primarily due to intensity gradients. This model suggests the following result:

$$\rho(X,Y) = \frac{\sigma^2_{XY}}{\sqrt{\sigma^2_{XX}\,\sigma^2_{YY}}} = \frac{\sigma^2_{xy} + 2\sigma^2_{st} + \sigma^2_{tt}}{\sigma^2_{ss} + 2\sigma^2_{st} + \sigma^2_{tt} + \sigma^2_{nn}} \quad \text{where} \qquad (2.23)$$

$$X = x + t + n_x \quad \& \quad Y = y + t + n_y \qquad \text{and}$$

$$\sigma_{ss}^2 = \sigma_{xx}^2 = \sigma_{yy}^2 \quad \& \quad \sigma_{st}^2 = \sigma_{xt}^2 = \sigma_{yt}^2 \quad \& \quad \langle n_x n_y \rangle = 0$$

Error signals were found not to affect experimental results greatly. So long as all correlation coefficients vary slowly, equation 2.23 appears well suited for comparison with the desired view expressed in equation 1.5. This assertion may be verified by inspection of actual interferograms, as for example in figure 2.2, for observable systematic errors.

Noise could be observed as a varying stitched pattern in cross-correlation interferograms, but was only significant in regions of low light intensity. Where signal is lacking, electronic noise dominates the measured video input. Such dark regions were kept well outside of the measurement region.

Texture, conversely, is visible in well lit regions, as a localized, bounded and somewhat recurrent phenomenon. With images of crisp focus, surface texture contribution per equation 2.23 is prone to great local variability. Such

Figure 2.2: Experimental cantilever interferogram

texture related error could be reduced by blurring the optics. In this manner, one should replace volatile σ_{st}^2 and σ_{tt}^2 by better behaved local averages.

Analysis of systematic cross-correlation errors thus leads to a positive conclusion. Interferograms of consistent quality may be generated quite readily. One finds that noise observed in real experimental fringe images corresponds almost entirely to statistical sampling error in the cross-correlation.

2.2 Image filters

One would ideally prefer to regard only raw unfiltered data while solving the inverse problem of this thesis. Such a goal cannot however be reconciled with another research aim: one requires sole reliance upon the single interferogram, for data determining the deformation field. One notes, with reference to equations 1.5 and 1.9, that the current treatment relies upon global rather than local interference effects.

The purist's solution to the noise problem lies nearer the traditional approach to *ESPI*. Variation of optical path allows for direct evaluation, pixel by pixel, of all quanitities expressed in formula 1.9. One moreover measures phase angle directly, rather than its cosine. Rather elegant phase unwinding schemes may be devised to remove angular wrap modulo 2π. The literature documents many investigations[DAIN 89][JONE 83] into related theoretical diversions.

The interferometer of the experiment at hand is simpler than the tradi-

tional one: no exotic optics are present for optical path variation in generation of interferograms. What this thesis investigates is how and to what extent one may recover the deformation field without such enhancements.

The method under study relies on less localized observation of interference. Formula 2.9 of the previous section shows that one expects noisy local signal behaviour. Image filtering of some type is clearly required, for identification and location of fringes in the background noise. Association of location and phase requires such noise reduction.

The impulsive thought toward weighting the *ESPI* data, through use of formula 2.18, must be exercised with some subtlety. Regions of high correlation and low noise correspond to relative phase angles $\Phi = N \cdot 2\pi$; The cross-correlation function, like $\cos\Phi$, should be stationary with respect to spatial variation along the gradient of Φ. More specifically, cross-correlation images exhibit least noise in high flat regions. Good discrimination of phase angle is difficult here. One finds, moreover, that even theoretically quiet points border on regions with too much experimental noise, for clear identification of local maxima.

Consequently, some image filtering is required for systematic identification of local maxima. The study of optimal filtering opens into an expansive field with many digressions.[JONE 83][SCHM 00] From practical experience with image processing, however, one finds a simple and consistent approach to similar imaging problems. The problem at hand requires a filter which reduces local noise sharply while preserving significant edges and gradients;

one might additionally like to smooth the data further. The industrial means to this end are simple and well known: apply a local median filter of some dimension, followed by a low pass or Gaussian blur filter. A comparison of the effects of such filters appears in Appendix A.2.

The filters employed in the present research were of precisely the industrial type. Images were treated via a rectangular (7×7) pixel median followed by a radius $\sigma = 5$ pixel circular Gaussian blur. Under such filtering, signals become quiescent, so that maxima and even minima may be clearly identified. Data about the maxima, however, do clearly remain most reliable. Examples of raw and filtered correlograms appear in section 5.2 as steps in the general image processing algorithm. The effect of the median filter in thresholding interferograms for maxima is discussed in subsection 4.2.3 below.

2.3 Data refits on raw and filtered images

According to the prediction of subsection 2.1.5, no advantage can be gained from use of data outside of cross-correlation maxima. Theory expects that inclusion of noisy data, in any phase/deformation fit, should in fact deteriorate fit quality. One may nonetheless attempt such a sub-optimal fit, and assign non-zero weight to data, which theory claims should be ignored. As an aside of the attempt, one may produce better scaled interferograms from raw and filtered cross-correlation images. One expects to produce cleaner images, but without improving fit precision.

A key problem with data located outside local maxima rests in the form of equation 1.5. Cross-correlations at local maxima should always value near unity. Away from local maxima however, variation of function $\kappa(\{i_A\},\{i_B\})$ becomes a material concern.

One may dispense with the variability of $\kappa(\{i_A\},\{i_B\})$ in purely empirical fashion. Since unmeasured cross-correlation errors may play a role, one rather considers the practical equation 1.9. Since functions $f(x,y)$ & $g(x,y)$ vary smoothly over the region of interest, one may approximate them with low order bivariate polynomials over x & y. In the present research, both functions were modelled as biquadratics. Biquadratics were chosen as the best compromise between real functional variation flexibility, in expressing function $\kappa(\{i_A\},\{i_B\})$, and the aim not to mimic real data fluctuations with the offset and scaling functions $f(x,y)$ & $g(x,y)$.

One makes use of a primary fit on maxima only. With apriori fit in hand, one simply forward models the deformation to yield the expected interferogram. For generation of a better scaled experimental interferogram, all that remains is to identify optimal biquadratic functions $f(x,y)$ & $g(x,y)$, which map the experimental cross-correlation to the theoretical result.

One may define the respective standard linear regression problem solely in terms of its χ- statistic as follows.[MATH 70] Let $Z(x,y)$ denote the theoretical interferogram and $A(x,y)$ refer to the raw or filtered experimental cross-correlation image. If angle brackets $\langle \ldots \rangle$ denote statistical expectation

in the ordinary sense, one minimizes the following:

$$\chi \;=\; \langle (Z(x,y) - f(x,y) - g(x,y) \cdot A(x,y))^2 \rangle \qquad (2.24)$$

While equation 2.24 does measure 18 parameters, the problem reduces simply to a standard linear regression of the form outlined in Appendix B.4.

After identifying optimal functions $f(x,y)$ & $g(x,y)$, one applies the corresponding transformation to the raw data, and acquires (fixed) experimental interferogram:

$$Y(x,y) \;=\; f(x,y) + g(x,y)\, A(x,y) \qquad (2.25)$$

One interprets $Y(x,y)$ directly as interferogram cosine data. Such a view assigns to each pixel a relative phase angle corresponding to its measurement in the raw interferogram. One chooses the branch of the inverse cosine as that nearest in value to the relative phase predicted by the primary maxima only fit. Functionally one has the following:

$$\phi(x,y) \;=\; \pm\cos^{-1}(Y(x,y)) + 2N\pi \quad \text{where} \qquad (2.26)$$

$$\pm \;\&\; N \quad \text{minimize } |\phi(x,y) - Z(x,y)|$$

The newly unwound phase image $\phi(x,y)$, given by equation 2.26, may be refit against the deformation inverse problem. As indicated in the foregoing, one expects unweighted secondary fits, of all data so generated, to deteriorate in quality from the primary maxima only fits. An experimental demonstration of this effect on a real cantilever will follow in section 5.3.

Chapter 3

Displacement fields

3.1 Basic interferometry

The mathematics of speckle interferometry are fairly simple to set out, the guiding physics[HECH 74] being no more complicated than Young's experiment. The formation of speckle patterns is due to simple wave interference, between the relative phases of a pair of exposures, from a pair of localized scattering sources. The problem of this thesis is hence most appropriately regarded from such a straight forward vantage.

Pearson's cross-correlation image patterns are no more than interference fringes. This has been established in chapters 1 and 2. Since general interference measures differential phase angle in sinusoidal format, one measures precisely the same in a cross-correlation interferogram. In examination of speckle patterns, one finds interference bands, with gradients pointing out the steepest rate of phase angle change. Curves of constant interference phase

amplitude, (such as maxima and minima,) correspond conversely to locii of constant phase.

The process of inverting fringe patterns into deformation fields may be regarded as a series of problems encountered in the imposition of no more complicated constraints than these. As will be shown in the following development, local phase gives an insufficient descriptor on the path to an inversion. Phase data alone do not contain enough information to identify a single causative deformation. In light of this revelation, moreover, one will find general inversion to be impossible. Curiously, even under constraint to conformal deformation fields, modes of deformation exist, which are not measurable in any inversion scheme.

Principally, the speckle interferometer is constituted, per figure 3.1, with a single laser, a beam splitter, a pair of scattering (source) targets to expand illumination, a deformable illuminated surface re-scattering laser light, and a recording camera system. Of these constituents, the only factors material for controlled production of speckle cross-correlation fringe patterns are the locations of the (source) targets, and the shape and orienation of the deformable surface re-scattering light. The experiment seems at least superficially identical to the classic Young's experiment: two coherent sources are projected onto a viewing surface.

The sources are however not as simple as idealized Young's double slits. Each scattering source is more properly seen as a localized (stochastic) collection of scattering sources, which from a distance may be realized as some

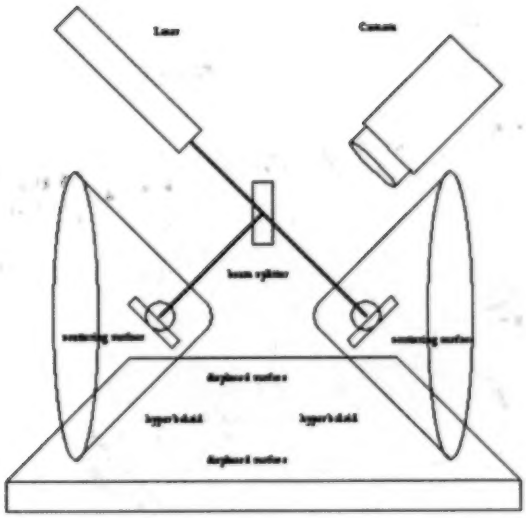


Figure 3.1: Speckle interferometer layout

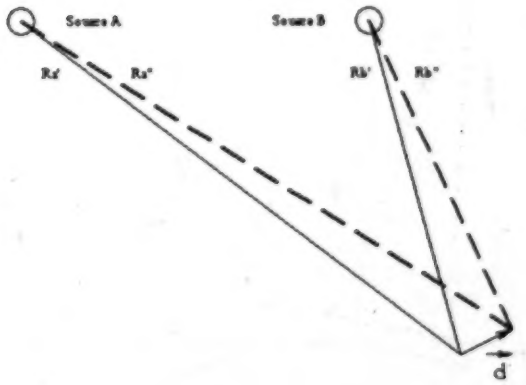


Figure 3.2: Effects of point displacement

high order multipole wave expansion. Such an expansion manifests globally in terms of vector harmonics, and locally in terms of linearized (plane wave) asymptotes. A discussion developing this point of view is found in Appendix A.5. The expansion itself drops out locally, as a stochastic phenomenon.

Pearson's cross-correlation in turn analyzes the randomizing statistics for relative phase information. The method renders clean virtual interference patterns like those from an idealized source pair. The result should not be so surprising, in that Young in his day could also not have produced a clean double slit source; he nonetheless was able to observe interference predicted by the perfect idealized model.[HECH 74]

If one views the scattering surfaces as sources in the style of Young's double slit experiment, the mathematics of fringe production become almost trivial. The pertinent feature of a point in space, determining the degree of constructive or destructive interference from a source pair, is just the difference in optical path from the respective coherent sources, as viewed from that point. From the path difference and the relative source phases, interference phase trivially follows.

Further predictions follow as well, by extrapolating locally from such a point in space into its neighbourhood. In moving away from a point along a direction altering beam path, interference phase variation occurs. Conversely, all points sharing the same path difference with respect to the source pair also share the same phase. This set of points corresponds to a unique hyperboloid through each point in space. All points on a particular hyper-

boloid share identical phase difference value. Hyperboloids corresponding to different values therefore cannot intersect. One may thus imagine a family of nested hyperboloids, parametrized by phase difference, as pictured in figure 3.1. It is in terms of such hyperboloids that the inversion problem is most clearly analyzed.

Most texts on the subject of speckle interferometry define the concept of a sensitivity vector at this point.[DAIN 89][JONE 83] From the view point of the hyperboloids, this vector is just the phase gradient, and runs perpendicular to the hyperboloids' surfaces. A graphic illustration the phase gradient vector appears in figure 3.2. This phase gradient may be expressed precisely as the following differential form:

$$\frac{\lambda}{2\pi}\Phi = r_A - r_B \tag{3.1}$$

$$\frac{\lambda}{2\pi}\tilde{d}\Phi = \tilde{d}r_A - \tilde{d}r_B \tag{3.2}$$

Here λ denotes the coherent laser wavelength, Φ gives the relative interference phase, and the other quantities are represented in figure 3.2. Linearized representation of changing phase (for tiny deformations) can be expressed in terms of the phase gradient:

$$\tilde{k}(x,y) = \frac{2\pi}{\lambda}\left(\frac{\tilde{r}_A}{r_A} - \frac{\tilde{r}_B}{r_B}\right) \tag{3.3}$$

$$\delta\Phi(x,y) \sim \tilde{k}(x,y)\cdot\vec{d}(x,y) \tag{3.4}$$

It is convenient to view computations for all deformations under study in terms of the (asymptotic) linearized representation of equation 3.4. The

result derives from the lowest order term in an asymptotic expansion of the radius scalar:

$$|\vec{r}+\vec{d}| \sim r + \hat{\vec{r}}\cdot\vec{d} + \frac{1}{2r}\vec{d}(1-\hat{\vec{r}}\otimes\hat{\vec{r}})\vec{d} \qquad (3.5)$$

$$\text{where} \quad r=|\vec{r}| \quad \text{and} \quad \hat{\vec{r}}=\frac{\vec{r}}{r}$$

With view to this expansion, linear approximation of relative phase is valid to fractional corrections on the order $O(\frac{d}{r})$. Since d ($\sim 10\mu m$) is at most a few times the laser wavelength and r ($\sim 10cm$) is of the order of the viewed surface dimension, one expects the approximation to be good to better than 5 significant figures. Notwithstanding linearization of relative phase, real surface deformation fields may vary more generally, as nonlinear functions of location over the viewed surface.

3.2 General displacement fields

One may easily imagine pathological deformations. Suppose, for example, that one may smoothly deform the surface in a precisely defined manner, to yield a definite measured fringe pattern. Even if the deformation and resulting fringe pattern are known precisely and without error, one finds that they cannot be invertibly linked. To show that this is the case, define a random mapping as follows. For each point on the viewed surface, choose a point sharing the same constant phase hyperboloid as the original point. Move the original surface point to the new location. This procedure may be applied individually to each point of the viewed surface. Subject to some physical

and mathematical constraints, one may thereby generate a pathological and possibly discontinuous deformation. If point deformations are small relative to observed speckle size and relative to image pixel dimension, neither image change nor loss of coherence would be observed. What should be observed, however, is an image with identically unaltered interferogram. One may thus generate a family of pathological deformations, with each member yielding identically the given interference fringe pattern.

The foregoing argument demonstrates that no true inversion is possible, which assigns to a given fringe pattern a unique deformation field. The resulting non-uniqueness holds even if apparatus dimensions are known precisely and the fringe pattern is measured without error. One is led, therefore, to question whether deformation families may nonetheless be identified with deformations of more restrictive descriptions. The class of random discontinuous deformations is certainly too general, as seen in the above argument. Restriction to experimentally realistic displacement fields will be shown to mitigate but not alleviate the uniqueness issue.

3.3 Stiff displacement fields

Realistic deformations are expected to be (almost) continuous over at least regions of the surface, outside of neighbourhoods about macroscopic fractures and fissures.[BHAT 86] In fact, one expects deformations to be fairly smooth on scales orders of magnitude larger than viewed pixel size or speckle

dimension. Subject to such restrictions, one may re-evaluate the issue of uniqueness.

For the present discussion one notes that general displacements include true deformations, as well as purely rigid surface motions. True deformations are in fact only defined up to a rigid displacement. The current study views displacement fields through their effective interferogram production. In this context, one regards the complete displacement field on the surface, and avoids confusing distinction between the terms. For the sake of simplicity, the term deformation will be understood as inclusive of all general dispacement fields modelled in the following development.

3.3.1 Rigid displacement fields

One considers here only the most restrictive of stiff deformations. Rigid transformations correspond identically to the full set of Killing vector fields in 3-dimensional space;[SCHU 80] they are most simply expressed as the set of combinations of translations and rotations.

Amongst rotations, one may single out an obvious anomaly. Rotation about the axis connecting the scattering sources demonstrates a unique symmetry. Since the axis of rotation is precisely the axis of all nested constant phase hyperboloids, each hyperboloid rotates and maps identically into itself. As a result, no surface alteration can be observed from interference phase change under such transformation: all such rotations belong to a family of general deformations with identical shared interferogram.

Physically, one may picture the anomalous rigid rotations as movements of a swing: if one imagines the surface as hanging suspended from a bar, connecting the scattering laser source pair, all microscopic *swing stage* motions to and fro share the same family interferogram.

Functional and linearized forms describing swing stage transformations may be expressed respectively as follows:

$$\vec{d}(x,y) = (\mathbf{R}(\hat{e}_{AB};\theta) - \mathbf{1})\,(\vec{r} - \vec{r}_{AB}) \tag{3.6}$$

$$\vec{d}(x,y) \sim \theta \cdot \vec{d}_{\text{swing}}(x,y) \quad \text{where}$$

$$\vec{d}_{\text{swing}}(x,y) = \vec{\hat{e}}_{AB} \times (\vec{r} - \vec{r}_{AB}) \quad \text{or} \tag{3.7}$$

$$\vec{d}_{\text{swing}}(x,y) = \frac{1}{2}\vec{\hat{e}}_{AB} \times (\vec{r}_A + \vec{r}_B)$$

Here A and B refer to the scattering sources and θ defines rotation angle in radians, and one defines geometric constants from the scattering source locations as follows:

$$\vec{\hat{e}}_{AB} = \frac{(\vec{A} - \vec{B})}{|\vec{A} - \vec{B}|} \quad \text{and}$$

$$\vec{r}_{AB} = \frac{1}{2}(\vec{A} + \vec{B})$$

For deformations of at most a few wavelengths magnitude, angles must be tiny, $O(\lambda/r)$ theoretically or $< 10^{-5}$ experimentally. Sinusoidal expansions

$$\sin(\theta) \sim \theta \cdot (1 - \frac{1}{6}\theta^2) \quad \text{and}$$

$$\cos(\theta) \sim 1 \cdot (1 - \frac{1}{2}\theta^2),$$

show second order corrections as negligible effects. The linearized form does clearly suffice in description of this singular effect.

The rigid *swing stage* rotation may be generalized by allowing the rotation angle θ to vary as a function over the deformation surface. The linearized form of this expression will be of importance in subsequent sections:

$$\begin{aligned} \vec{d}(x,y) &= (\mathbf{R}(\hat{e}_{AB};\theta(x,y)) - \mathbf{1})\,(\vec{r} - \vec{r}_{AB}) \\ \vec{d}(x,y) &\sim \theta(x,y) \cdot \vec{d}_{\text{swing}}(x,y) \end{aligned} \tag{3.8}$$

Regardless of the nature of the function $\theta(x,y)$, so long as its magnitude is small in terms similar to those of the rigid transformation of formula 3.6, no effect will be noted in the corresponding interferogram. If one requires that displacement fields on the surface vary slowly and smoothly, then function $\theta(x,y)$ might parametrize a slowly and smoothly varying invisible component of the total variation function.

The question arises whether other deformations exist, invisible to interferometric observation, and if so, how they might be described. Both questions are answered by generating exactly such a function. With help of the cross-product, a *singular* deformation field appears perpendicular to the sensitivity vector and to the *swing stage* deformation field:

$$\begin{aligned} \vec{d}_{\text{sing}}(x,y) &= \frac{\lambda}{\pi}\,\vec{k}(x,y) \times \vec{d}_{\text{swing}}(x,y) \quad \text{or} \\ \vec{d}_{\text{sing}}(x,y) &= \vec{\tilde{e}}_{AB}\,[\,(r_A - r_B)\,(1 - \tilde{\tilde{e}}_A \cdot \vec{\tilde{e}}_B)\,] \\ &- (\vec{r}_A + \vec{r}_B)\,[\,\tilde{\tilde{e}}_{AB} \cdot (\vec{\tilde{e}}_A - \vec{\tilde{e}}_B)\,] \end{aligned} \tag{3.9}$$

The *singular* deformation field clearly has far more complexity than does the *swing stage* rotation. Close examination shows that no choice of function

$\phi(x,y)$ leaves a simple expression for the functional product $\phi(x,y)\cdot\vec{d}_{\text{sing}}(x,y)$; square root and inverse square root forms do not simplify into finite order polynomials. One makes use of this observation in following sections.

One may now make some fundamental observations. First and foremost, the vector set $\{\vec{k}, \vec{d}_{\text{swing}}, \vec{d}_{\text{sing}}\}$ forms an orthogonal basis for 3-dimensional deformations. As such, any microscopic deformation field may be uniquely expressed in terms of functions $\psi(x,y)$, $\theta(x,y)$ and $\phi(x,y)$ as follows:

$$\vec{d}(x,y) \sim \psi(x,y)\cdot\vec{k}(x,y) + \theta(x,y)\cdot\vec{d}_{\text{swing}}(x,y) + \phi(x,y)\cdot\vec{d}_{\text{sing}}(x,y) \quad (3.10)$$

Of the contributions, only the first, being parallel to the sensitivity vector, can be measured. Sensitivity vector $\vec{k}(x,y)$ however, expressed in formula 3.3, fails to include realistic deformations such as the Killing vectors,[SCHU 80] and realistic strains which are their smooth local *gauge* variants.[BHAT 86] The first term describes instead yet another deformation field of peculiar curvature.

One should rather like to target measurement at realistic deformation fields. The generalized *gauged* Killing vector terms are clearly amongst the contributions one must address. These contributions include, as base terms, rigid motions which have vanishing strain components. Smooth variation of parameters generating rigid motions gives the associated *gauge* terms corresponding to smoothly varying strain fields. Such deformations are the most common subject of experimental measurement.

As *gauge* functions vary more flexibly, most realistic smooth deformation

functions may be expressed within their span. Smooth and slowly varying deformations of such description, by construction, have significant components tangential to the span of $\{\vec{k}, \vec{d}_{\text{swing}}\}$; realistic fields are unlikely however to show the curvature of $\vec{d}_{\text{sing}}(x, y)$. By convenient coincidence, therefore, most realistic deformations may be locally measured by restriction of displacement fields to constant rigid transformations perturbed by slowly varying strain fields.

Expansion of deformations in terms of *gauged* Killing vector fields seems a most expedient approach. One at once includes all terms necessary for extraction of macroscopic strains, but still limits curvature of measured deformation fields. In such an approach one may fit realistic *stiff* deformation fields to interferogram images.

3.3.2 Polynomial displacement fields

As initial forray into study of polynomial deformations, Killing vectors near the deformation surface may be expressed as polynomial deformations. Inspection of the *swing stage* deformation of formula 3.7 shows this special rotation to be a monomial in cartesian coordinates. The general Killing vector may similarly be expressed, as included by the general monomial in cartesian coordinates, as follows:

$$\begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} + \begin{bmatrix} s_{xx} \\ s_{xy} + r_z \\ -r_y \end{bmatrix} \cdot x + \begin{bmatrix} s_{xy} - r_z \\ s_{yy} \\ +r_x \end{bmatrix} \cdot y \qquad (3.11)$$

Here, polynomial expansion is with regard to cartesian coordinates (x, y)

over the deformation surface. The inferred z coordinate runs normal to the surface. The Killing vector displacement field $\vec{d}$ is expressed in terms of translation $\vec{t}$, and linearized (antisymmetric) rotation $\vec{r}$. Remaining terms comprise the (symmetric) tangential or in-plane dilation $\vec{\vec{s}}$; this tensor expresses all non-rigid (non-Killing) components at monomial order. One observes further, that constant translation terms fill out the 0^{th} order polynomial, while rotations and tangential dilations complete the 1^{st} order terms.

In light of the monomial expansion, one notes that the concept of the *gauged transformation* simplifies, in polynomial terms, to simple functional multiplication. In terms of order statistics, it becomes sensible to deal directly with polynomials of given order, since polynomial functions are, by construction, simple multiplicative functions of the linearized Killing vector fields. With such indirect expansion of *gauged* Killing vectors, further questions of overconstraint and underconstraint may be answered simply in terms of maximal polynomial order.

In demonstration of such an approach, consider as prime example a general polynomial deformation field of order N, with applied interferometrically invisible perturbation:

$$\vec{d}(x,y) \;=\; \vec{P}_N(x,y) \,+\, Q_{N-1}(x,y)\cdot\vec{d}_{\text{swing}}(x,y) \tag{3.12}$$

The bivariate polynomial deformation $\vec{P}_N(x,y)$ is observable. Scalar polynomial $Q_{N-1}(x,y)$, parametrizing the invisible part may be varied arbitrarily to generate a family of maximally N^{th} order polynomial deformation fields

sharing the same interferogram.

Conversely, one may devise a consistent reduction method which maps a family of deformations $\vec{d}(x,y)$,expressed via equation 3.12, to a single result. Intuitively, polynomial division seems the sensible approach. Vector field $\vec{Q}_{N-1}(x,y)$ cannot however be reduced to a scalar. One may nonetheless carry out order by order optimization in some sense.

In writing this thesis, the author chose to minimize residual coefficient magnitude (in quadrature) successively at each order from highest to lowest:

$$\begin{aligned} \text{for} \quad \vec{P}_N(x,y) &= \sum_{n=0}^{n=N} \sum_{m=0}^{m=n} \sum_{i=1}^{i=3} P_{nmi}\, x^{(n-m)} y^m \hat{e}_i \\ \text{minimize} \quad \chi_n &= \frac{1}{3n} \sum_{m,i} P_{nmi}^2 \quad \text{for} \quad n = N, \ldots, 1 \end{aligned} \tag{3.13}$$

The prescribed algorithm projects an entire family expressed in formula 3.12 onto a single polynomial deformation. Thus, even though unique inversion fails, measurable differences in deformation fields may be identified directly from measured polynomial coefficients.

The polynomial approach does however suffer from an immediate drawback: degrees of freedom expand steeply as polynomial order increases. Illustrating this point, one may count the coefficients S_N, defining scalar polynomials of order N, over 2 independent variables x and y as:

$$S_N = \frac{1}{2}(N+1)(N+2) \tag{3.14}$$

The number C_N of vector polynomial degrees of freedom, the rank R_N of the observable fit space, and the dimension K_N of the singular (kernel) space

spanned by *gauged swing stage* transformations, respectively, may be deduced from a fit as follows:

$$C_N = \frac{3}{2}(N+1)(N+2), \tag{3.15}$$

$$R_N = (N+1)(N+3) \quad \text{and} \tag{3.16}$$

$$K_N = \frac{1}{2}N(N+1). \tag{3.17}$$

The kernel dimension K_N grows quickly from 0, asymptotically (ie. $N \to \infty$) accounting for 1/3 of all degrees of freedom. Low order values appear displayed in table 3.1. Clearly, limitation of deformations and models to low order fits is well advised. Higher order fits are limited asymptotically, by the form of equation 3.10, to one of the total three functional degrees of freedom.

The next problem is to set up a practical fit mechanism. The goal is to fit actual polynomial deformations to phase angle data from given observed interferograms. A pair of approaches to the problem present themselves. Between these methods, the less direct turns out to be the better choice in terms of both versatility and stability.

N	C_N	R_N	K_N
0	3	3	0
1	9	8	1
2	18	15	3
3	30	24	6
4	45	35	10

Table 3.1: Polynomial coefficients, fit rank and kernel size

The obvious method involves apriori elimination of all non-observable *gauged swing stage* terms from a polynomial basis set, followed by a fit and finally by expression of the result as a polynomial. This method, however, has some undesirable attributes. One immediately increases the problem's complexity, while failing to deal with a fundamental problem: any optimization over a large parameter basis lacks stability. Calculation on a finite precision machine compounds fragilities inherent to the data.

Elimination, order by order, of overdefining basis elements may be accomplished in theory by several methods. Problems arise in practice, from the dimension of the fit. Because one is fitting many parameters, fit stability depends strongly on choice of region, balance of data and interferometer geometry. Fits of this type generally require a great deal of fine tuning.

On completion, a meaningful solution is hoped for, but success is by no means guaranteed. Upon achieving any solution, results remain suspect. One has yet to question how close the fit was to being singular. In practice one would prefer to gauge the quality of a solution, and if possible, identify singularities directly, checking their contribution to the result.

From the standpoint of utility, a method should yield reliable results in a sensible fashion, and gauge the quality of a solution, possibly even diagnosing the causes of potential algorithm failure. If possible, one would like to carry out interferometric deformation measurement in the following order: first, generate a realistic "stiff" (low order polynomial) deformation field, and second, deal with non-observable components.

Expression of a final result in invariant terms is the ideal goal. Simply stated, successive solutions for the same deformation should yield identical solutions. Results should not vary by a polynomial offsets in the nature of equation 3.12.

Surprisingly, such an algorithm does actually exist. Solution of the problem on the above stated terms requires a generalized form of linear regression analysis: simple matrix inversion solutions do not suffice for the problem considered here. The more general method of *singular value decomposition (SVD)* does however yield a textbook solution.[PRES 92]

3.4 Singular value decomposition

The method of *singular value decomposition (SVD)* has several advantages which may be demonstrated by example.[PRES 92] First and foremost, one must express a (non-negative semi-definite) quadratic form as χ- statistic for linear regression analysis:

$$\chi \;=\; \langle\, (\, \phi(x,y) \;-\; \phi_0 \;-\; \tilde{k}(x,y) \cdot \vec{P}_N(x,y)\,)^2 \,\rangle \qquad (3.18)$$

Field $\phi(x,y)$ corresponds to phase angle as measured in an interferogram image. Vector field $\vec{k}(x,y)$ is just the spatially varying sensitivity vector of equation 3.3. Optimization is carried out, respecting as parameters base offset ϕ_0 and coefficients of polynomial deformation $\vec{P}_N(x,y)$. Angle brackets $\langle \ldots \rangle$ denote statistical expectation in the ordinary sense.

One optimizes such a quadratic form, in the normal fashion, as laid out

in Appendix B.4. The greatest complication here, from a computational standpoint, involves ordering parameters in sensible fashion. With view to sequential storage and processing in the C language,[KERN 88] a most sensible order may be prescribed as follows.

The polynomial deformation field may be expanded as a sum, over a vector basis $\langle \hat{e}_1, \hat{e}_2, \hat{e}_3 \rangle = \langle \hat{e}_x, \hat{e}_y, \hat{e}_z \rangle$, as well as powers of x and y:

$$\vec{P}_N(x,y) = \sum_{n=0}^{n=N} \sum_{m=0}^{m=n} \sum_{i=1}^{i=3} P_{nmi}\, x^{(n-m)} y^m\, \hat{e}_i \tag{3.19}$$

The index i runs fastest and n runs slowest in evaluation of this sum. The actual deformation fit functions may now be expressed as a set, in terms of polynomial displacement coefficients P_{nmi}, and sensitivity vector field (equation 3.3) components $k_i(x,y)$, with ordering imposed by sum 3.19:

$$f_{nmi}(x,y) = x^{(n-m)} y^m \cdot k_i(x,y) \tag{3.20}$$

Finally one has only to express functions as a contiguous linear vector:

$$f_k(x,y) = \begin{cases} 1 & k = 0 \\ f_{nmi}(x,y) & k = 1 + \frac{3}{2}n(n+1) + 3m + i \\ & i \in \{1,2,3\} \;\&\; m \in \{0,\ldots,n\} \end{cases} \tag{3.21}$$

The resulting curvature matrix $\boldsymbol{C}$ and data vector $\vec{v}$ may be expressed in the normal fashion of linear regression:

$$C_{ij} = \langle\, f_i(x,y) \cdot f_j(x,y) \,\rangle \quad \text{and} \quad v_i = \langle\, \phi(x,y) f_i(x,y) \,\rangle \tag{3.22}$$

The problem cannot however be solved in the normal fashion: curvature matrix $\boldsymbol{C}$ cannot be inverted. More subtlety is required here, but simple

matrix methods do suffice for almost magical recovery of all stability, with full analysis of the remnant instability space.

Symmetry, $C_{ij} = C_{ji}$, shows that $\boldsymbol{C}$ may be diagonalized. The kernel of $\boldsymbol{C}$ is spanned by eigenvectors with vanishing eigenvalues. All *gauged swing stage* transformations are included here. Data instabilities may well add to the kernel's rank. In computational terms, any sufficiently small eigenvalue should raise concern.

The problem inherently contains a generating source for functions of precisely such description. The dramatic example arises as fit polynomial order is increased; existence of a polynomial of order at most N, which may be approximated by $\psi(x,y) \cdot \vec{d}_{\text{sing}}(x,y)$ for some nontrivial function $\psi(x,y)$, will correspond to an eigenvector with small eigenvalue.

Fortunately, *singular value decomposition* allows for identification of such problems, and experiment in fact shows that under reasonable interferometer arrangements, such terms do not become significant until fits of at least 4^{th} order are attempted over entire images. Localized solutions of 1^{st} order are stable even on relatively minor regions of an image.

Solution of the *singular value decomposition* requires application of a final inference. The optimization solution vector solves, as in the case of normal linear regression, $\vec{v} = \boldsymbol{C}\vec{a}$. Any vector $\vec{x}$ in the kernel of $\boldsymbol{C}$ thus satisfies:

$$\forall \vec{x} \quad \boldsymbol{C}\vec{x} = \vec{0} \quad \longrightarrow \quad \vec{x} \cdot \vec{v} = \tilde{x}\boldsymbol{C}\vec{a} = \tilde{a}\boldsymbol{C}\vec{x} = 0 \tag{3.23}$$

Plainly stated, vector $\vec{v}$ has no component tangential to the kernel of $\boldsymbol{C}$.

One may therefore limit the original curvature matrix $\boldsymbol{C}$ and vector $\vec{v}$ by omitting the kernel. Solution of the optimization problem reduces to performing inversion on the nonsingular remnant. One thus finds solution vector $\vec{a}$ which satisfies generalized linear regression equation:

$$\forall \vec{x} \quad \boldsymbol{C}\vec{x} = \vec{0} \quad \longrightarrow \quad \vec{v} = \boldsymbol{C}(\vec{a} + \vec{x}) \tag{3.24}$$

The process of *singular value decomposition (SVD)* outlined in the foregoing demonstrates the theoretical utility required for inversion of interferograms into displacement fields. Practical implementation of such methods however involves more complexities. Setup and inversion of the experimental problem of this thesis cannot be accomplished directly in finite precision math from the derivation presented in this section. Some computational refinements are required if one wishes avoid recourse to extended precision arithmetic. Calculations must be arranged so that floating point sums and differences avoid information loss. The interested reader is referred to Appendix C.1.

3.5 Source geometry adjustment

One may analyze interferometer precision from the perspective of effects in *SVD* linear regression. Errors may be classified by origin into a pair of categories. The first category entails variation in experimental geometry, due to physical effects during acquisition of a video image pair. Induced errors affect only observations taken from a particular pair. Causative factors here include

vibrational displacements and thermal expansion or contraction. The second category derives systematically from errors in specification of experimental parameters. These errors derive from apparatus measurement imprecision. Geometric deviation from assumed values induces errors in all interferometer observations. Both error classes relate measurably with geometric interferogram perturbations.

3.5.1 Geometry variation between images

The *singular value decomposition* method of equation 3.18 contains a hidden geometry correction as follows. By presumption, the constant phase offset term ϕ_0, in the base *SVD* problem, serves primarily for identification of the central maximum fringe, where phase $\phi(x, y)$ vanishes. The solution should return, within error, a whole wave number. In practice, however, any slight motion of a scattering source target or laser aim can cause a non-integral phase shift.

The mathematics of source and laser displacement extend the sensitivity vector derivation of equations 3.1 through 3.4. In the general case at hand, one varies the viewed surface, and as well, varies laser and scattering source locations. Continuing to label the scattering sources by their location vectors $\vec{A}$ & $\vec{B}$ and simplifies the respective laser sources to point emanations from $\vec{L}_A$ & $\vec{L}_B$, one may express for any point $\vec{r}$ on the deformation surface, the relative phase through the entire optical path:

$$\frac{\lambda}{2\pi}\Phi = (|\vec{A} - \vec{L}_A| + |\vec{r} - \vec{A}|) - (|\vec{B} - \vec{L}_B| + |\vec{r} - \vec{B}|) \quad (3.25)$$

$$\begin{aligned}
\frac{\lambda}{2\pi}\delta\Phi &\sim [\delta\vec{r} - \frac{1}{2}\delta(\vec{A}+\vec{B})]\cdot\vec{\Delta}_r - \frac{1}{2}\delta(\vec{A}-\vec{B})\cdot\vec{\Delta}_s \\
&+ \delta(\vec{A}-\vec{L}_A)\cdot\vec{\Delta}_A - \delta(\vec{B}-\vec{L}_B)\cdot\vec{\Delta}_B; \\
\vec{\Delta}_r &= \frac{\vec{r}-\vec{A}}{|\vec{r}-\vec{A}|} - \frac{\vec{r}-\vec{B}}{|\vec{r}-\vec{B}|} \\
\vec{\Delta}_s &= \frac{\vec{r}-\vec{A}}{|\vec{r}-\vec{A}|} + \frac{\vec{r}-\vec{B}}{|\vec{r}-\vec{B}|} \\
\vec{\Delta}_A &= \frac{\vec{A}-\vec{L}_A}{|\vec{A}-\vec{L}_A|} \\
\vec{\Delta}_B &= \frac{\vec{B}-\vec{L}_B}{|\vec{B}-\vec{L}_B|}
\end{aligned} \tag{3.26}$$

Several new and interesting consequences appear in such expression of the geometric variation. Most immediately, however, one recognizes in the first term the familiar sensitivity vector $\vec{k} = \frac{2\pi}{\lambda}\vec{\Delta}_r$. Other contributions derive from new vector elements.

As one expects, the $\vec{\Delta}_r$ term shows symmetric motions of the sources to be indistinguishable from converse global vector translations of the entire deformation surface. One might also suppose that antisymmetric motions of the sources be indistinguishable from complementary global rotations and dilations. Expansion of related transformations shows limited agreement with this hypothesis.

Rigid rotations about the centroid of the scattering sources may be expanded as inclusions in the $\vec{\Delta}_s$ term of equation 3.26. One may verify such a claim by noting the following identity for cross-product rotation generators:

$$[(\vec{A}-\vec{B})\times\vec{R}]\cdot\vec{\Delta}_s = [\vec{R}\times(2\vec{r}-\vec{A}-\vec{B})]\cdot\vec{\Delta}_r \tag{3.27}$$

The related dilation, however, conceals a decoherent effect. Identities for translation and rotation terms, described above, derive from the existence of global rigid transformations (Killing vectors) moving the entire apparatus. The dilation term does not preserve experimental geometry, in that lengths measured in laser wavelengths do not remain fixed. One may, conversely, view the dilation in terms of an equivalent perturbation of laser wavelength, followed by rigid deformations of the types described in the preceding. In this view the adjustment to laser wavelength is as follows:[JONE 83]

$$\frac{\delta\lambda_D}{\lambda} = \frac{\delta_{\|}|\vec{A}-\vec{B}|}{|\vec{A}-\vec{B}|} \tag{3.28}$$

Since optical paths vary with location $\vec{r}$ on the deformation surface, the $\vec{\Delta}_s$ dilation term does indeed introduce decoherent relative phase errors.

One finally notes that the last pair of terms shows coherence entirely independent of surface location $\vec{r}$. One may interpret the net contribution as a deviant global perturbation of the observed phase offset:

$$\phi_{\text{dev}} \sim \delta(\vec{A}-\vec{L}_A)\cdot\vec{\Delta}_A - \delta(\vec{B}-\vec{L}_B)\cdot\vec{\Delta}_B \tag{3.29}$$

One may recognize such variation of phase offset as an included function in the *SVD* fit of equation 3.18. Microscopic motions of the optics with respect to scattering spot location are thereby already reasonably treated.

From the standpoint of experimental design, one infers, in view of equation 3.26, that the effect of the $\vec{\Delta}_s$ dilation term should be mitigated by mounting sources rigidly onto the same structural base. Other changes in

setup, including motion of that structural base with regard to the viewed deformation surface, as well as motion in the optics (laser mount, splitting reflector, etc.) are well circumscribed by the $\vec{\Delta}_r$ translations, $\vec{\Delta}_\theta$ rotations, and $\vec{\Delta}_A$ & $\vec{\Delta}_B$ phase offset terms.

In light of mounting constraints and the real distances involved, one must rely on the predictability of these effects, in that none of the causative variations can be entirely ruled out.

3.5.2 Geometry variation from specifications

As demonstrated above, most variations in experimental geometry, which occur during the course of an interferometer observation, are already theoretically addressed in the *SVD* fit. One might reasonably ask whether systematic offsets in source geometry may be dealt with directly by an interferometric technique. Such a method would allow for fine tuning of the interferometer's geometric paramaters. One could thus account for variations which remained constant throughout sets of experiments.

Such measurement tuning does in fact become possible by extension of the *SVD* linear regression of the last section. By extension of the the basis function set to include gradient terms along the lines of the second order correction of equation 3.5, one may write a χ- statistic expanding that of equation 3.18 in the following form:

$$\chi = \quad \langle (\tag{3.30}$$
$$\phi(x,y) - \Phi - (\tilde{\hat{r}}_A - \tilde{\hat{r}}_B) \cdot \vec{f}(x,y)$$
$$- \tfrac{1}{r_A} \tilde{f}(x,y) \, (\mathbf{1} - \vec{\hat{r}}_A \otimes \tilde{\hat{r}}_A) \, \vec{a}$$
$$- \tfrac{1}{r_B} \tilde{f}(x,y) \, (\mathbf{1} - \vec{\hat{r}}_B \otimes \tilde{\hat{r}}_B) \, \vec{b}$$
$$)^2 \rangle$$

The non-negative semi-definite quantity of equation 3.30 would be equivalent to that of equation 3.18 but for the augmentation of the pair of vector parameters $\vec{a}$ & $\vec{b}$ and their respective functions. The new least squares problem expressed in equation 3.30 is not however linear. The last lines of equation 3.30 contain quadratic terms in fit parameters.

One may however express the problem as an iteration which is linear at each step, by eliminating the nonlinearity as follows:

$$\chi_k = \quad \langle (\tag{3.31}$$
$$\phi(x,y) - \Phi_k - (\tilde{\hat{r}}_A - \tilde{\hat{r}}_B) \cdot \vec{f}_k(x,y)$$
$$- \tfrac{1}{r_A} \tilde{f}_{k-1}(x,y) \, (\mathbf{1} - \vec{\hat{r}}_A \otimes \tilde{\hat{r}}_A) \, \vec{a}_k$$
$$- \tfrac{1}{r_B} \tilde{f}_{k-1}(x,y) \, (\mathbf{1} - \vec{\hat{r}}_B \otimes \tilde{\hat{r}}_B) \, \vec{b}_k$$
$$)^2 \rangle$$

With natural choice $\vec{f}_0(x,y) = \vec{0}$, one may acquire a solution sequence, by iterative minimization for a succession of ordered sets. The algorithm obtains a solution sequence of the form: $\{\langle \chi_k, \Phi_k, \vec{f}_k, \vec{a}_k, \vec{b}_k \rangle : k \in \mathcal{W}\}$.

One would like to interpret new solution vectors $\vec{a}$ & $\vec{b}$ as corrections to the respective source locations $\vec{A}$ & $\vec{B}$:

$$\begin{aligned} \frac{\vec{r}-(\vec{A}-\vec{a})}{|\vec{r}-(\vec{A}-\vec{a})|} &\sim \left|\vec{\hat{r}}_A+\tfrac{1}{r_A}\right|(1-\vec{\hat{r}}_A\otimes\vec{\hat{r}}_A)\,\vec{a} \\ \frac{\vec{r}-(\vec{B}+\vec{b})}{|\vec{r}-(\vec{B}+\vec{b})|} &\sim \left|\vec{\hat{r}}_B-\tfrac{1}{r_B}\right|(1-\vec{\hat{r}}_B\otimes\vec{\hat{r}}_B)\,\vec{b} \end{aligned} \tag{3.32}$$

Note the adjustments $\vec{A}' = \vec{A} - \vec{a}$ & $\vec{B}' = \vec{B} + \vec{b}$ indicated by the asymptotic forms. Interpretation of the new parameters $\vec{a}$ & $\vec{b}$ thus becomes transparent: they do indeed amount to linearized corrections for source locations.

Through application of the prescribed adjustments to source locations, one should in theory improve base measurements for the sources, and improve interferometer precision. Such a measurement technique shows the promise of utility, in that base measurement of scattering source location involves the awkwardness of measurement around mounted structures.

Part II

Experiments

Chapter 4

Image processing

4.1 Intensities

Laser speckle intensity distributions are easily measured from video images. Consistency of theoretical distributions expressed in equations 2.2 and A.39 may thus be tested against experimental observation. Raw image intensity histograms (with adjustment for only video offset) need only to be viewed. Histogram distribution characteristics give a good test for agreement between theory and experiment: Comparison of peak locations and magnitudes, as well as low and high intensity drop-off characteristics, may be compared. All comparisons may further be made in native relative intensity measure. Theory makes no use of absolute intensity values, with distributions depending solely on relative statistics for observable parameters. These tests are purely statistical and physically trivial.

The limiting case where speckle size at least eclipses pixel dimension was

Figure 4.1: Speckle pattern: single scattering source, $\sim 45^o$ incidence

well beyond the resolution of the actual interferometer. Equivalent speckle patterns were generated by projecting a red laser diode through translucent tape onto a white wall. A 35mm lens was unscrewed on a *CCD* camera to simulate microscope magnification, for generation of images resembling figure 4.1. In that image, the field of view has dimension of roughly $2mm$, and was illuminated by a source approximately $1.5m$ away.

Experimental intensity histograms, in figures 4.2 and 4.3, show good agreement with theoretical curves for relative intensity. One may compare respective parameters, via equation 2.2, for polarization degree. Polarized scatter (figure 4.2 with $\mu = 15.8$ & $\sigma = 12.0$) gave principal intensities $i_\perp = 10.9$ & $i_\top = 4.9$. The nonpolarized case (figure 4.3 with $\mu = 30.0$ & $\sigma = 21.6$) gave corresponding intensities $i_\perp = 17.9$ & $i_\top = 12.1$. Polarization degree may indeed be deduced from intensity histogram parameters.

Parameter measurements indicate that, even without external polarization filters, laser light scatter seems partially polarized. While this may be

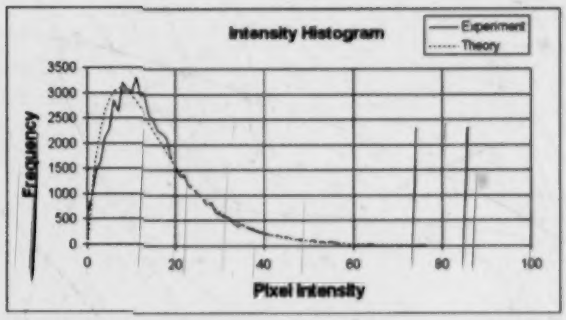


Figure 4.2: Polarized intensity vs. eqn 2.2 with $\mu = 15.8$ & $\sigma = 12.0$

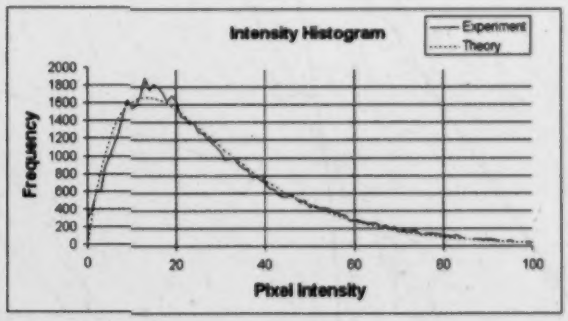


Figure 4.3: Nonpolarized intensity vs. eqn 2.2 with $\mu = 30.0$ & $\sigma = 21.6$

true, investigation shows theoretical curves to be insensitive under small variation of $i_\perp$ & $i_\top$. One infers that polarization degree cannot be measured with any real precision from such histograms.

The result corresponding to a badly blurred view of a similar optical setup is shown in figure 4.4. Statistics $\mu = 108.9$ & $\sigma = 36.4$ generate an asymptotic gamma function with $\alpha = 15.8$ & $\beta = 9.14$. The gamma fit is

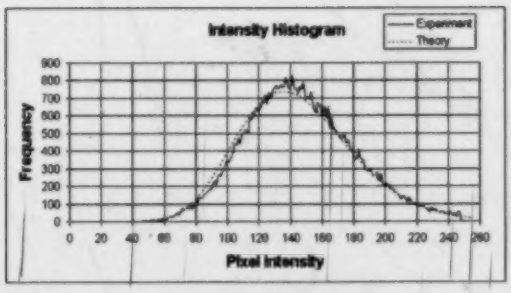


Figure 4.4: Blurred intensity vs. eqn 2.1 with $\mu = 108.9$ & $\sigma = 36.4$

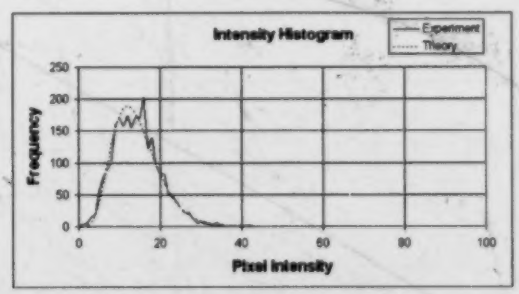


Figure 4.5: Experiment: single source vs. eqn 2.1 with $\mu = 14.23$ & $\sigma = 5.63$

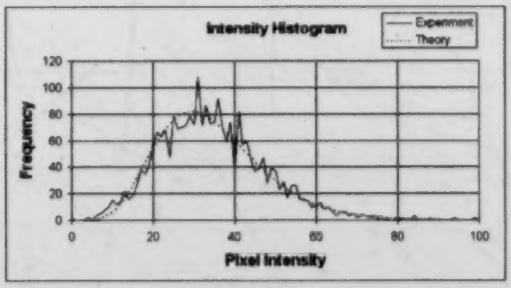


Figure 4.6: Experiment: dual sources vs. eqn 2.1 with $\mu = 34.4$ & $\sigma = 13.0$

quite reasonable and the elevation of the α parameter agrees well with theory.

The experimental interferometer was examined under illumination by a single scattering source, and under illumination by both. Identical 50×50 pixel^2 regions of the deformation surface yielded intensity distributions shown in figures 4.5 and 4.6, respectively. Figure 4.5 with $\mu = 14.23$ & $\sigma = 5.63$ has $\alpha = 6.4$ & $\beta = 2.2$, whereas figure 4.6 with $\mu = 34.4$ & $\sigma = 13.0$ has $\alpha = 7.1$ & $\beta = 4.9$. The asymptotic gamma function hypothesis of equation A.39 agrees well with figures 4.5 and 4.6.

While pixel dimension, polarization degree and angle of incidence affect precise details of the result, theory does predict that the gamma density parameter α value for the single source should be lower than that for dual sources. The respective values $\alpha = 6.4$ (single source) and $\alpha = 7.1$ (dual sources) agree with theory on this point.

4.2 Filters

One may readily produce experimental evidence for the theoretical claims, from subsections 2.1.3, 2.1.4, 2.1.5 and 2.1.6, regarding cross-correlation errors. First, the prediction of equation 2.23 should be tested, that blurred images might produce cross-correlations of sharper contrast. Second, the distribution of the linearized error term of equation 2.9 should be examined for its appearance in real experimental data. Third, cross-correlation variability has to be demonstrated near local maxima, in order to justify use of

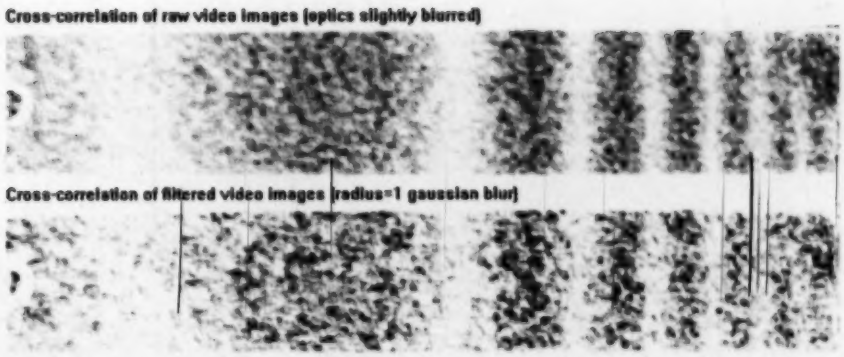

Cross-correlation of raw video images (optics slightly blurred)
Cross-correlation of filtered video images (radius=1 gaussian blur)

Figure 4.7: Blur effect: video blur enhances cross-correlation contrast

image filtering in identification of local maxima.

4.2.1 Video blur effect

The effect of real optical blurring may be simulated by applying a Gaussian blur filter to a pair of raw speckle video images. With subsequent cross-correlation of respective raw and filtered image pairs, results may be compared, as displayed in figure 4.7.

The cross-correlation of blurred images shows contrast improvement over the cross-correlation of the original raw images. It also shows increased noise levels in regions of low cross-correlation, as predicted by equation 2.9. Despite increased noise, the improved contrast does however offer improved resolution, for isolation of local maxima from noisy neighbouring data. Since local maxima contain all the real information in interferogram fringes, blurring

does improve data recovery.

One obtains better contrast yet when optics rather than digital images are blurred. Propagation of errors from signal digitization and processing roundoff is mitigated. All experimental interferograms used for fits in section 5.3 were generated with a slight but deliberate optical blur.

4.2.2 Cross-correlation sampling errors

Equation 2.1.3 makes a suspect claim: the linearized error might not apply with complete generality for all cross-correlated samples. One may check its verisimilitude by testing it against the entire set of cantilever data. Define random variable Q by the ratio between actual and linearized variance described in equation 2.11. Expressly one may write:

$$Q = \frac{N}{2} \frac{\sigma_{\rho\rho}^2}{(1-\rho^2)^2} \tag{4.1}$$

The distribution of Q was tested against all the cantilever data, with data binned vertically for each horizontal pixel count. All interferogram images were used, with only edge data being omitted from the enumeration. The resulting histogram appears in figure 4.8.

The mean value $\bar{Q} =\sim 80\%$ compares reasonably well with unity, indicating agreement with the linearized error estimate. One has also, for the case with total pixel overlap, good agreement with the minimum dimension estimate of equation 2.16. Experimental $\beta = 2.2$ agrees well with theoretical χ^2-value $\beta = 2$; standard deviations $\sigma_Q =$ 28% &30% (experiment & the-

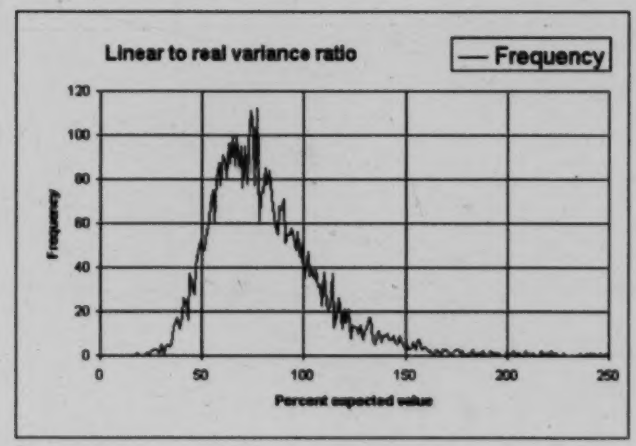


Figure 4.8: Experimental distribution for $Q = \frac{N}{2} \frac{\sigma^2_{\hat{e}\hat{e}}}{(1-\rho^2)^2}$ (85 pixel samples)

ory) agree correspondingly well. The remainder of data tabulated in table 4.1 were enumerated in much the same manner. Data samples were however selected sparsely, avoiding overlap cross-correlation in the sense of equation

Sample	ν_{thy}	ν_{exp}	β_{exp}	μ_{exp}	σ_{thy}	σ_{exp}	$\% - \sigma_{\text{err}}$
4	3	3.4	1.7	79	65	60	-7
6	5	5.0	2.0	82	52	52	0
8	7	6.1	2.3	84	45	48	+7
10	9	7.4	2.4	85	40	44	+11
12	11	8.4	2.6	85	36	42	+15
14	13	9.5	2.7	86	34	39	+17
16	15	10.3	2.9	85	31	37	+21
85 (min ν)	17	15.4	2.2	82	28	30	+5
85 (trace ν)	25	15.4	3.2	82	23	30	+27
85 (max ν)	84	15.4	10.9	82	13	30	+133

Table 4.1: Experimental distribution for $Q = \frac{N}{2} \frac{\sigma^2_{\hat{e}\hat{e}}}{(1-\rho^2)^2}$ (various samples)

2.13. Results show good agreement with the prediction of subsection 2.1.3, that Q follow a χ^2-distribution, and with degrees of freedom at approximately the expected number. Standard deviation estimates σ agree, as one expects, even better than sample ν & β parameters.

With all the underlying approximations, one must conclude that theory and experiment agree far better than one might have expected. Equation 2.10 indicates that the approximation should lack accuracy with 5×5 cross-correlation samples. Further precision loss derives from assigning vertical lines a single relative phase value. One must clearly regard with skepticism the minimum dimension approximation of equation 2.16. As a result, experiment falls closer to theoretical approximation than one might have hoped.

Due to good agreement with experimental observation, one might find a more practical purpose, for distributional analysis of quantity Q. Doubts regarding clear identification of pixel subpopulations, corresponding to fixed relative phase angles, may be resolved by observing the distribution's tendency to spread. Thus Q, ν and β give measure, in some sense, to the quality of discrimination between subpopulations, corresponding to differing cross-correlation values. In the case of table 4.1, for example, the identification of vertical lines as single phase bins lacks perfection; some of the observed deviation from theoretical values does stem from lack of clean phase discrimination.

Since figure 4.8 and table 4.1 agree well with the linearized error estimate of subsection 2.1.3, one infers that the linear estimate of equation 2.9

applies to experimental cross-correlation interferograms. One may moreover conclude that the optimal assignment of statistical weights as calculated in equation 2.20 comports with experimental statistics; optimal deformation fits result when only data from local maxima are utilized.

4.2.3 Location of maxima

Based on the foregoing, one may streamline enumeration of phase information. Identification of local fringe maxima is theoretically simple when viewed through equations 1.2 and 1.5. While cross-correlation values are elsewhere subject to smooth variability function $\kappa(\{i_A\}, \{i_B\})$, local maxima always correspond identically to $\rho = 1$. Experimental observation agrees on this point. Local maxima may be identified to constrained regions by applying an identical optimal global threshold to all cross-correlation images, as depicted in figure 4.9, with above threshold maxima marked in black.

Signal noise about local maxima derives from subpopulations, with cross-correlations near unity, neighbouring on those maxima. One would like to filter out the noisy boundary regions leaving only local extrema. Note that any filter stripping images for these central maxima must necessarily induce some statistical bias. While better filters do exist, a simple method with acceptable bias presents itself.

The median filter simply and naturally selects spatially localized dominant populations. The corresponding bias respects tight bounds, as demonstrated in Appendix A.2. The image may be thresholded as before, with a

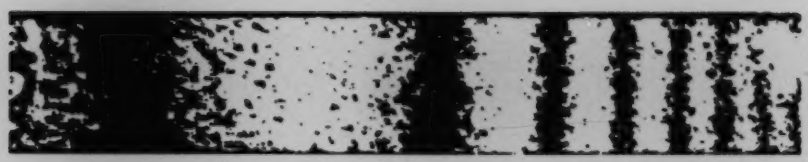


Figure 4.9: Cross-correlation threshold for local maxima (raw data)

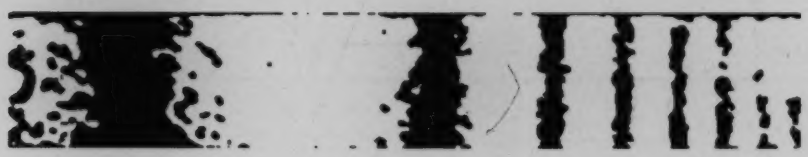


Figure 4.10: Cross-correlation threshold for local maxima (median data)

single optimal global threshold, yielding far cleaner results, as depicted in figure 4.10, with above threshold maxima again marked in black.

One may further smooth images with a blur filter, as done in the next section of this research. Blurring however further compromises amplitudes near local maxima. The advantage of a single global threshold is lost as the cost of further smoothing.

The theoretical value for such an optimal threshold may be evaluated in this rare instance. Derivation and discussion of the empirical effects of this result will follow in section 5.3. Experimentally it was found that intuitive notions of optimality suffice in setting a global threshold.

Chapter 5

Deformations

5.1 Model data

The algorithm development stage of the research involved coding and testing of the complex *singular value decomposition (SVD)* methods outlined in subsection 3.4 and Appendix C.1. The process required generation and testing of software against model data. Test simulations included rigid translations and rotations, as well as polynomial deformations. Results included matching input and output interferograms; more precisely, input and output deformations were compared, and the number and nature of (singular) kernel vectors were examined for consistency with theoretical claims.

Interferometer parameters were identified and structured for computational utility. Imposition of an orderly approach allowed for testing the more exotic source geometry corrections described in section 3.5. Algorithms for geometry adjustment were tested with simulated rigid transformations.

5.1.1 Rigid translations

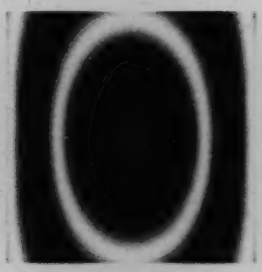


Figure 5.1: Translation: $\delta x = 2\mu m$

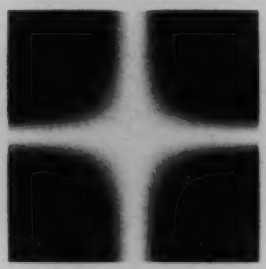


Figure 5.2: Translation: $\delta y = 2\mu m$

Figure 5.3: Translation: $\delta z = 2\mu m$

5.1.2 Rigid rotations

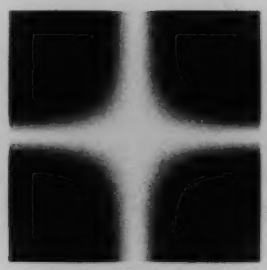


Figure 5.4: Rotation: $\theta_x = 20n$radians

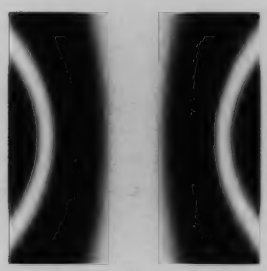


Figure 5.5: Rotation: $\theta_y = 20n$radians

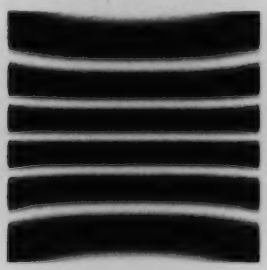


Figure 5.6: Rotation: $\theta_z = 20n$radians

5.1.3 Rigid simulations

Figures 5.1 through 5.6 were generated by forward modelling polynomial deformations, viewed through a particularly simple interferometer arrangement. The arrangement was chosen principally to illustrate the effect of the *swing stage* deformation. Note the similarity of figures 5.2 and 5.4. The images and all underlying phase information are in fact identical. One may moreover scale both deformations by any small constant, to yield yet another pair of indistinguishable interferograms. The effect is of such a fundamental nature. One may nonetheless study the inversion problem via *singular value decomposition* as presented in section 3.4.

In order to preserve generality and utility of the method, a complete set of parameters was constructed to define the interferometer. The general interferometer was defined through a text interferometer *(IFM)* data file. The file was formatted as shown in figure 5.7. An *IFM* file contains laser wavelength (mm), six scattering source (spot) coordinates (mm), eleven camera calibration parameters, two image pixel spatial dimensions (mm), and input video array width and height (pixels) and (output) perspective corrected array width and height (pixels). All distances are assumed to share consistent reference frame and units. In the example, the reference frame is left handed and all measurements are in millimeters.

Using the exact interferometer parameter file of figure 5.7, the limits of *singular value decomposition* techniques were readily tested. Each of the

```
0.0006328     //laser wavelength
+75.00000     //spot A        -  x-position
0.0000000     //spot A        -  y-position
+100.0000     //spot A        -  z-position
-75.00000     //spot B        -  x-position
0.0000000     //spot B        -  y-position
+100.0000     //spot B        -  z-position
136.428321    //afRcam[0]
70.798688     //afRcam[1]
188.383081    //afRcam[2]
-1.165873     //afEcam[0]
-0.617763     //afEcam[1]
-2.096616     //afEcam[2]
319.500000    //Pixel         -  x-centre
239.500000    //Pixel         -  y-centre
1192.765563   //Pixel         -  x-magnification
1192.765563   //Pixel         -  y-magnification
2.652759      //fChi
0.400000      //Grid x-gap    -  out pixel x-dimension
0.400000      //Grid y-gap    -  out pixel y-dimension
512.0000      //Grid width    -  out array in pixels
512.0000      //Grid height   -  out array in pixels
640.0000      //Image width   -  in array in pixels
480.0000      //Image height  -  in array in pixels
```

Figure 5.7: Model data interferometer parameters in *IFM* file format

images in figures 5.1 through 5.6 were subjected to the general inversion technique with first and second order polynomial deformations. With input data rounded to single precision floating point, fits returned parameters identical to the generating inputs, with errors at $\sim 10^{-5}$, near the limit of single precision floating point accuracy.

The *swing stage* deformation was eliminated from fit solutions, to yield "reduced equivalent" deformations, by the algorithm of equation 3.13. Vectors in the kernel of the curvature tensor, corresponding to the null eigenvectors, were tested for consistency with the *gauged swing stage* hypothesis by dividing them out by the same algorithm.

A fictitious inner product was devised to test the *residue* outside the *gauged swing stage* function span. In order to keep the result independent of fit region and data set, one could make use of interferometer design order parameters to define a test inner product for polynomials (in the form of equation 3.13) as follows:

$$\langle \vec{P}_N, \vec{Q}_N \rangle = \sum_{n=0}^{n=N} \sum_{m=0}^{m=n} \sum_{i=1}^{i=3} |\vec{s}|^n P_{nmi} Q_{nmi} \qquad (5.1)$$

$$\text{where } \vec{s} = \frac{(\vec{A}-\vec{B}) \times (\vec{A}+\vec{B})}{2|\vec{A}-\vec{B}|}$$

The residue for a fit was defined as the average value, for vectors $\vec{K}_N$ spanning the fit kernel, of the quantity $\langle \vec{k}_N, \vec{k}_N \rangle / \langle \vec{K}_N, \vec{K}_N \rangle$, where $\vec{k}_N$ is the remainder when the division procedure 3.13 is applied to $\vec{K}_N$.

All fits on theoretical rigid deformations met expectations, with fit residues at the limit of double precision floating point accuracy ($\sim 10^{-10}$), indicating

stable algorithm success, and as described in the foregoing, floating point precision accuracy ($\sim 10^{-5}$) results on floating point precision data. Parameter agreement from the test fits indicated that the algorithm was stable and workable on extremely clean data.

One could simulate experimental conditions by adding speckly noise. Two sets of tests were conducted to investigate the effect of data noise. Both tests led to roughly the same conclusion: Algorithm stablility depends strongly on clean data.

A simple noise simulation amounted to rounding input data to 8-bit fractional accuracy (to the nearest 256^{th}). First order polynomial fits remained reliable to $\sim 10^{-3}$ whereas second order fits deteriorated sharply in consistency. While parameter matches deteriorated, interferograms generated by original and fit parameters continued to be in good agreement.

The second simulation amounted to a test of the phase-unwinding refit algorithm per equation 2.26, simulating experimental refit conditions as follows. The phase offset term ϕ_0 of equation 3.18 was intentionally altered by 1/4 wave. The second test gave results similar to those of the first. First order fit consistency survived the noise, while second order fit consistency deteriorated.

Results agree fairly well with expectations. The algorithm under current study requires a source of clean input data. Noise gives rise to instability of the *SVD* algorithm, especially as the number of fit parameters increases.

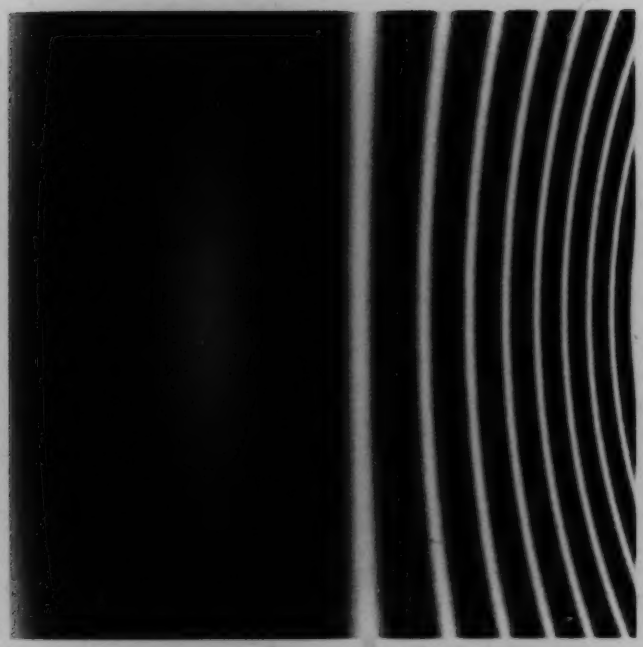


Figure 5.8: Cantilever interferogram (40g) – extrapolated theoretical

5.1.4 Polynomial deformations

More realistic deformations could be investigated as well. With regard to experimental tests in the next section, one chose to view the theoretical deflection of a simple cantilever, fixed at one end and loaded at the other. The interferogram in figure 5.8 results, from precisely the measured experimental interferometer geometry, described in the *IFM* data file of figure 5.9.

The non-ideal cantilever grid shape, with real dimensions $140mm \times 25mm$ has width:height ratio of 700 : 125, which presents a serious problem. All

```
0.0006328     //laser wavelength
+132.00000    //spot A      -  x-position
-2.50000      //spot A      -  y-position
+73.00000     //spot A      -  z-position
-125.00000    //spot B      -  x-position
-2.50000      //spot B      -  y-position
+79.00000     //spot B      -  z-position
136.428321    //afRcam[0]
70.798688     //afRcam[1]
188.383081    //afRcam[2]
-1.165873     //afEcam[0]
-0.617763     //afEcam[1]
-2.096616     //afEcam[2]
319.500000    //Pixel       -  x-centre
239.500000    //Pixel       -  y-centre
1192.765563   //Pixel       -  x-magnification
1192.765563   //Pixel       -  y-magnification
2.652759      //fChi
0.200000      //Grid x-gap  -  out pixel x-dimension
0.200000      //Grid y-gap  -  out pixel y-dimension
700.0000      //Grid width  -  out array in pixels
125.0000      //Grid height -  out array in pixels
640.0000      //Image width -  in array in pixels
480.0000      //Image height -  in array in pixels
```

Figure 5.9: Cantilever interferometer parameters in *IFM* file format

attempts failed to perform the general *SVD* fit for a cubic order deformation in the form of equation 3.19 on the peculiar experimental grid shape.

By extrapolating the grid to square (700:700) shape, however, as displayed in figure 5.8 the inverse problem yielded readily to the theoretical method, as before, with single precision floating point input data. The parameter match from the fit is practically perfect, as tabulated in table 5.1. Coefficients in

Coefficient	Theory	*SVD* Fit	Difference
$10^3 \cdot \phi_0/2\pi$	-0.003	0.030	0.033
$10^0 \cdot P_{000}$	0.000	0.000	0.000
$10^0 \cdot P_{001}$	0.000	0.000	0.000
$10^0 \cdot P_{002}$	-4.964	-4.965	-0.001
$10^2 \cdot P_{100}$	0.000	0.000	0.000
$10^2 \cdot P_{101}$	0.000	0.000	0.000
$10^2 \cdot P_{102}$	-10.197	-10.197	0.000
$10^2 \cdot P_{110}$	0.000	0.000	0.000
$10^2 \cdot P_{111}$	0.000	0.000	0.000
$10^2 \cdot P_{112}$	0.000	0.000	0.000
$10^4 \cdot P_{200}$	0.000	0.000	0.000
$10^4 \cdot P_{201}$	0.000	0.000	0.000
$10^4 \cdot P_{202}$	-3.615	-3.615	0.000
$10^4 \cdot P_{210}$	0.000	0.000	0.000
$10^4 \cdot P_{211}$	0.000	0.000	0.000
$10^4 \cdot P_{212}$	0.000	0.000	0.000
$10^4 \cdot P_{220}$	0.000	0.000	0.000
$10^4 \cdot P_{221}$	0.000	0.000	0.000
$10^4 \cdot P_{222}$	0.000	0.000	0.000
$10^6 \cdot P_{300}$	0.000	0.000	0.000
$10^6 \cdot P_{301}$	0.000	0.000	0.000
$10^6 \cdot P_{302}$	1.724	1.724	0.000
$10^6 \cdot P_{310}$	0.000	0.000	0.000
$10^6 \cdot P_{311}$	0.000	0.000	0.000
$10^6 \cdot P_{312}$	0.000	0.000	0.000
$10^6 \cdot P_{320}$	0.000	0.000	0.000
$10^6 \cdot P_{321}$	0.000	0.000	0.000
$10^6 \cdot P_{322}$	0.000	0.000	0.000
$10^6 \cdot P_{330}$	0.000	0.000	0.000
$10^6 \cdot P_{331}$	0.000	0.000	0.000
$10^6 \cdot P_{332}$	0.000	0.000	0.000

Table 5.1: Cantilever *SVD* fit results – extrapolated grid

the table are corrected for the elimination of the 6 *gauged swing stage* deformations via the algorithm of equation 3.13. By experimental design, none of the *swing stage* deformations interfere with observation of an expected deformation mode.

One may draw a pair of conclusions from the theoretical polynomial fit. First, one could in theory solve the general inverse problem, from interferogram to deformation field, for deformations up to cubic order. This would be possible despite the 6 dimensional *SVD* kernel, subject to a few provisos. The data grid must have dimensions of sufficient size, and the interference has to be clean of noise to ($\sim 10^{-5}$) single precision floating point accuracy. In light of these constraints a second more practical conclusion follows. The real cantilever deformation should fail to yield to the full cubic fit, due to higher order contributions in a real deformation, due to the limitation of grid height, and due to statistical noise in the real data.

5.1.5 Source geometry algorithms

Adjustments to interferometer geometry outlined in section 3.5 were examined in detail with simulated and experimental data. The possibility of deviant non-integral wave-number phase offsets, in the sense of equation 3.29, was reflected consistently in all algorithm tests, on both simulations and real data. Deformation polynomial solutions consistently included deviant phase offset terms. Source geometry tuning, in the sense of equation 3.31, was tested principally on simulations, and only in passing on real data.

In light of the limitations posed by *SVD* methods under experimental conditions, poor results might be expected from the source geometry adjustment algorithm, as prescribed in equation 3.31. Surprisingly, quite the opposite is found. Example tests indicated that the method returned consistent and stable results on clean simulated data, and with a single interferogram sufficing for performance of the tuning procedure.

Rigid motions of a deformation surface are relatively easy to arrange; the technique outlined in equation 3.31 may be employed with polynomial deformations of first order. The first nontrivial terms in the solution sequences for vectors $\vec{a}$ & $\vec{b}$ might be interpreted as geometric corrections to scattering source location, in the sense of equation 3.32. One may apply the corrections directly to the interferometer parameter settings outlined in the interferometer's *IFM* file.

Error tolerance of the procedure was quite remarkable, with spot position deviations of $10mm$ magnitude vanishing rapidly in the iteration. Interferometer parameters reflected in figure 5.7 were perturbed from those used in the model simulation. Deformations corresponding to interferograms of figures 5.1 through 5.3 were tested as interferometer geometry tuning data, to recover the correct interferometer source geometry. Iteration results are displayed in table 5.2.

Subsequent attempts were made to tune experimental interferometer geometry for the cantilever. The experimental tests on real empirical data proved less successful. Scattering source locations did in fact converge with

Tuning (mm scale) via translation $\delta x = 2\mu m$ (figure 5.1)						
iteration k	δx_A	δy_A	δz_A	δx_B	δy_B	δz_B
0	+10.00	-10.00	+10.00	-10.00	+10.00	-10.00
1	-2.39	+1.00	+4.76	-2.82	+3.09	-1.31
2	+0.05	+0.10	+0.16	+0.03	-0.10	+0.36
3	0.00	0.00	0.00	0.00	0.00	0.00

Tuning (mm scale) via translation $\delta y = 2\mu m$ (figure 5.2)						
iteration k	δx_A	δy_A	δz_A	δx_B	δy_B	δz_B
0	+10.00	-10.00	+10.00	-10.00	+10.00	-10.00
1	-6.45	+3.63	+0.07	-4.40	+1.72	-1.91
2	-0.37	+0.12	+0.33	0.00	+0.04	+0.20
3	-0.09	-0.02	0.00	-0.05	-0.01	-0.05

Tuning (mm scale) via translation $\delta z = 2\mu m$ (figure 5.3)						
iteration$k$	δx_A	δy_A	δz_A	δx_B	δy_B	δz_B
0	+10.00	-10.00	+10.00	-10.00	+10.00	-10.00
1	+9.48	+2.00	+10.06	-6.95	+1.23	+0.99
2	+0.76	-0.18	-2.26	+1.21	-0.05	+0.76
3	+0.02	0.00	-0.04	+0.03	0.00	0.00

Table 5.2: Geometry tuning (mm scale) by rigid translation interferograms

stability to new values, but with impossibly large vector displacements:

$$\delta\vec{A} = (-15.5, +1.7, -18.9)\ mm$$

$$\delta\vec{B} = (+27.8, -7.0, -14.4)\ mm$$

Displacement magnitudes from measured source positions A and B were clearly well outside of any reasonably expected (millimeter scale) error range. One concludes, that limitations posed by the real cantilever data ruled out source geometry tuning, in this experimental instance.

5.2 Experimental data analysis

Evaluation of a deformation field may be carried out via a well defined procedure. Since much of the analysis involves image processing, the algorithm may be outlined best by demonstration. As will become evident, some minimal apriori knowledge of the nature of a deformation field is required for generation of a precise solution. The nature of this knowledge is limited though to recognition of phase gradients in an interferogram. Outside of this step, processing is primarily of a mechanical nature.

One begins processing with a camera view of the experimental setup as appears in figure 5.10. The camera view is subject to perspective correction and as well to optical aberrations. One may correct the view through use of a calibration grid, via polynomial correction or as was the case in this experiment, via proprietary camera calibration. The calibrated image may be resampled bicubically to give a view such as shown in figure 5.11.

The principal advantage of the proprietary method *(VisionSmart™ 2000)* is that camera location and orientation derive as parameters from the optimization. In this manner one may measure apriori camera location (optical centre) and compare with optimal values as a check on measurements. Since all measurements must be rationalized into a single (left handed) coordinate system, one finds this feature quite advantageous.

For the current examination, one may compare figures 5.12 and 5.13. The washer indicates the coordinate origin; the 5g weight demarks the intended

Figure 5.10: Camera calibration (raw)

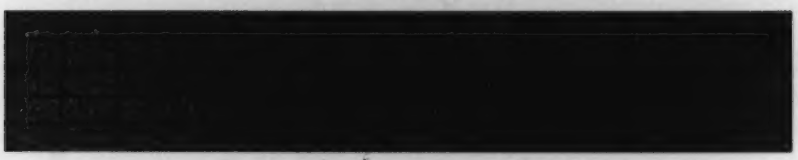


Figure 5.11: Camera calibration (perspective corrected)

Figure 5.12: View orientation (raw image)

Figure 5.13: View orientation (perspective corrected)

top left image corner. Allen keys point out scattering source locations, and the bolt points back to the camera. The oriented image, figure 5.13, deems right and down as respective positive x and y directions, with upward positive z direction completing the requisite left-handed coordinate system. The check on camera location agreed with rough measurements to within 2mm.

Data analysis requires a pair of image captures via video frame grabber. The first capture comprises a reference image. The second capture measures the deformation. Such an ordered image pair is shown in figures 5.14 and 5.15. Pearson's cross-correlation between the image pair yields a resulting interferogram depicted in figure 5.16.

The cross-correlation is then corrected for perspective via bicubic interpolation. The result is shown in figure 5.17. One could also perspective correct reference and video images before cross-correlation. Since edge effects and resampling cause slight statistical variation in observed cross-correlation, raw image data (figures 5.14 and 5.15 for example) should be used, rather than perspective corrected resampled images, for the purpose of cross-correlation image computation.

Simple image processing filters are applied to the cross-correlation image. The problem of smoothing an image with variable noise levels has a standard solution: a 7×7 median is applied to remove spike noise. A subsequent radius 5 Gaussian blur removes local statistical variation. These filter parameters were fixed for the entire data set. The resulting image is shown in figure 5.18. While contrast is reduced, observable interference fringes do appear to

Figure 5.14: Cantilever reference image (0g)

Figure 5.15: Cantilever video image (15g)

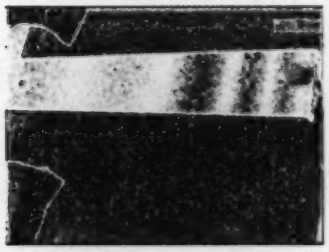


Figure 5.16: Cross-correlation (15g) with (5×5) samples

Figure 5.17: Cross-correlation (perspective corrected)

Figure 5.18: After 7×7 median and radius 5.0 Gaussian

-	-	0	0	0	0	0	0	+	+	+	0	0	0	0	0	0	-	-
-	-	0	0	0	0	0	0	+	+	+	0	0	0	0	0	0	-	-
-	-	0	0	0	0	0	0	+	+	+	0	0	0	0	0	0	-	-

Table 5.3: Cantilever extremum filter

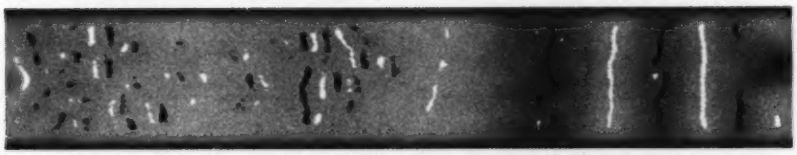


Figure 5.19: Identification of local extrema

vary far more continuously.

Fringe patterns of the type shown in image 5.18 may be searched for local extrema by applying a nonlinear filter. For the cantilever, one expects vertical banding and searches for maxima and minima via a filter of the type shown in table 5.3. The filter works as follows: when overlaid on the image, the central pixel is flagged as a local maximum whenever any pixel overlaid by '+' is above average and every pixel overlaid by '-' is below average on that region; if exactly the opposite is true, the central pixel is flagged as a local minimum; in either case pixels overlaid by '0' are not checked for parity with regard to the average. Application of this filter to figure 5.18 yields local interference extrema as displayed in figure 5.19. These results do show good agreement with theory, in that locations found for local minima are far more poorly defined than are those corresponding to local maxima.

The researcher at this point must intervene, by isolating the well defined local maxima, and rejecting noisy regions as exist in the left half of figure 5.19. Finally, a gradient is chosen to rank the fringe order, as is shown by the almost vertical grey-level bands, in figures 5.20 and 5.21. The latter restricts the former to the cantilever's right half image.

One next performs the primary data fit, on fringe maxima flagged in figures 5.20 and 5.21. Fits run on the data set fall roughly into two classes: cantilever specific fits and general *singular value decompositions*. Results of the respective fits are then forward modelled, to simulate interferograms corresponding to optimal fit functions, as depicted in images 5.22 and 5.23.

Figure 5.20: Gradient by grey shade on local maxima (full image)

Figure 5.21: Gradient by grey shade on local maxima (right half image)

Figure 5.22: Theoretical interferogram: primary (maxima only) cantilever fit

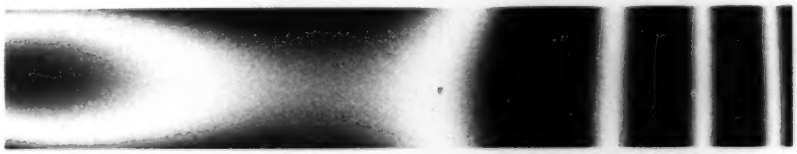


Figure 5.23: Theoretical interferogram: primary (maxima only) *SVD* fit

Figure 5.24: Raw result with biquadratic scale and offset to cantilever fit

Figure 5.25: Filtered result with biquadratic scale and offset to cantilever fit

Real experimental fits were performed on full and reduced (right half image) data sets; only results from the full image sets are shown here.

Interferogram deformation fits are, at this point, performed on data from interference maxima data only. Theory dictates that such fits give optimal fit results, as predicted in subsection 2.1.5. One may demonstrate the phenomenon, by utilizing larger regions of images for phase information. A biquadratic functional fit is first performed, along lines described in section 2.3. This fit of observed to theoretical interferograms applies biquadratically varying scale and offset functions to the experimental interferogram, and effectively maps cross-correlations into phase difference cosines. The process may be applied to either to raw or filtered images. The respective results from the cantilever model fit are shown in figures 5.24 and 5.25.

The biquadratic transformation imposes a fixed scale and offset on oscil-

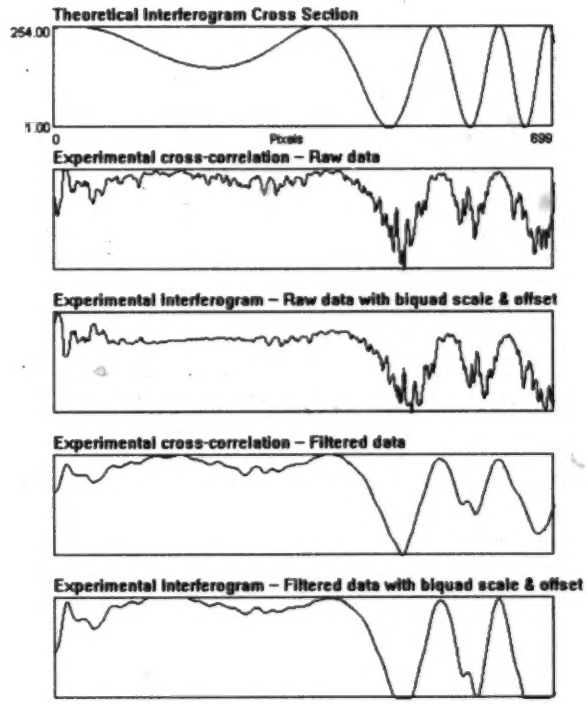


Figure 5.26: Cross-section correction: raw data give true cosine values

latory data. Consistency of results in images (figures 5.24 and 5.25) and in central cross-sections (figure 5.26) comports with cosine wave interpretation.

One next unwinds the effective cosine image data, by utilizing the primary (maxima only) fit, to fix the branch of the inverse cosine. The phase value assigned to an image pixel is defined, per equation 2.26, by the inverse cosine nearest in value to the primary fit's theoretical value. One generates a new phase data image by repeating the procedure over the entire interferogram.

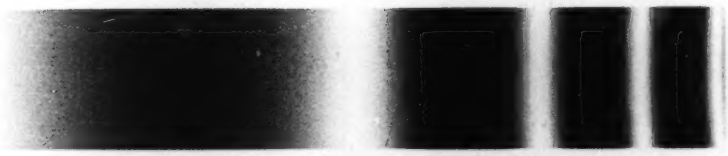


Figure 5.27: Theoretical secondary cantilever fit (raw phase unwound)

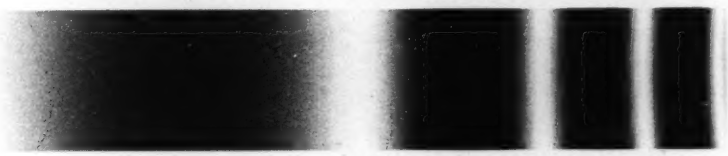


Figure 5.28: Theoretical secondary cantilever fit (filtered phase unwound)

A resulting new phase image may in turn be subjected to another similar fitting procedure. Forward simulations from such iterated fits are shown in figures 5.27 and 5.28.

One might in theory repeat the procedure. In practice, however, fit quality deteriorates with each successive attempt. Interferogram data away from local maxima tend to corrupt the fit sample, with noise orders of magnitude above that found near constructive fringe extrema. The interference noise profile may be observed directly, in the raw experimental interferogram cross-section of figure 5.26. This view agrees with perceptions of noise in raw cross-correlations, as in figure 5.17, and with observed distributions of minima and maxima, as found in figure 5.19. Noise levels near local minima are clearly orders of magnitude above those about local maxima.

5.3 Cantilever deformation

Initial attempts at fitting cantilever data to a general deformation met with failure. The general cubic polynomial deformation was immediately ruled out, for reasons outlined in subsection 5.1.4. The obvious compromise, of fitting a general quadratic deformation near the end of the cantilever failed, for much the same reasons. A subsequent attempt, to solve for cubic deformation limited solely to vertical displacements, failed as well.

The reason for such failures could be clearly identified as follows. Limitation of image data to a small rectangular region, in the form of a cantilever, greatly compromises resolution of functional differences. The short span along the y-axis is particularly harmful. Retreat to solution of a lower order equation, near the free end of the cantilever, actually worsens the situation, by creating a similar insensitivity along the x-axis.

If data precision were not so critical a factor, one might have been able to overcome this difficulty: the fit curvature matrix, and thereby the error matrix as well, are calculated entirely from independent variables; matrix precision is not an experimental issue, but rather a purely computational matter. Measured data needed only to have precision, sufficient to survive a single matrix multiplication, per equation 3.22. Computational stability of $\vec{a} = E\vec{v}$ was ruled out, however, by error in data vector $\vec{v}$ alone.

Despite lacking precision required for reliability of *SVD* algorithms, one could nevertheless produce suspicious first order polynomial fits to the

interferograms. Results were tabulated for fits covering most of the cantilever surface, in table 5.4, and for fits covering only data near the free end, in table 5.5. Constant displacement fits for near end data were also attempted, with results in table 5.6. Such zero-th order polynomial fits are inherently free of singular values.

One would not have expected sensible results from a low order fit on the entire surface of the cantilever. Low order end fits, however, either with the 1^{st} order *SVD* method or the 0^{th} order linear regression might have done better with less noise. As an interesting consequence of the suspicious fits, one could produce model interferogram simulations for the unrealistic results. Theoretical and experimental interferograms were found to match with remarkable consistency, both for near end fits and for the *SVD* fits over most of the cantilever surface.

Subsequently, the theoretical cantilever equation was fit against the same experimental interferograms. These fits were of the familiar sort, where a single parameter model deformation simulates interferograms, and parameter variation gives a best fit respecting the particular theoretical model. In the research at hand, the fits actually contained, as second free parameter, a global offset of the type described in equation 3.29. Fits of experimental data, against the theoretical boundary value problem, gave meaningful and consistent agreement between predicted and observed deviations.

Results from example fits appear respectively in tables 5.7 (full cantilever), and 5.8 (end data only). Since weight applied to the cantilever's

All Data: SVD (order=1) raw maxima fit ($\{x,y,z\} \rightarrow \mu m$)							
M_{app} (g)	δx_{end}^{cant}	δx_{end}^{exp}	δy_{end}^{exp}	δz_{end}^{exp}	$2\pi\,\phi_{dev}$	χ_{fit}/ν	res($\emptyset$)
10.0	3.3	95.2	4.5	-433.8	0.09	7.3E-07	1.6E-14
15.0	5.0	8.6	0.7	15.4	0.50	3.3E-04	9.8E-16
20.0	6.6	-1.3	2.0	15.6	-0.40	1.3E-03	4.4E-16
25.0	8.3	-6.7	0.4	14.2	0.21	1.1E-03	5.5E-16
30.0	10.0	-11.6	0.6	15.9	-0.35	2.0E-03	4.7E-16
35.0	11.6	-10.4	-0.3	16.4	0.31	2.6E-03	8.7E-16
40.0	13.3	-4.2	0.3	24.3	-0.03	2.5E-03	1.6E-16
μ						1.4E-03	2.7E-15

All Data: SVD (order=1) raw phase unwind refit ($\{x,y,z\} \rightarrow \mu m$)							
M_{app} (g)	δx_{end}^{cant}	δx_{end}^{exp}	δy_{end}^{exp}	δz_{end}^{exp}	$2\pi\,\phi_{dev}$	χ_{fit}/ν	res($\emptyset$)
10.0	3.3	95.7	4.3	-434.9	0.43	1.8E-02	2.5E-16
15.0	5.0	7.9	0.9	16.2	-0.05	4.6E-03	2.5E-16
20.0	6.6	0.9	1.2	16.5	-0.31	5.6E-03	2.5E-16
25.0	8.3	-2.1	1.2	18.5	0.15	8.1E-03	2.5E-16
30.0	10.0	-8.0	-0.2	18.4	0.05	8.2E-03	2.5E-16
35.0	11.6	-9.5	0.2	16.9	-0.46	1.0E-02	2.5E-16
40.0	13.3	-5.9	0.2	23.9	0.25	8.7E-03	2.5E-16
μ						9.1E-03	2.5E-16

All Data: SVD (order=1) filtered phase unwind refit ($\{x,y,z\} \rightarrow \mu m$)							
M_{app} (g)	δx_{end}^{cant}	δx_{end}^{exp}	δy_{end}^{exp}	δz_{end}^{exp}	$2\pi\,\phi_{dev}$	χ_{fit}/ν	res($\emptyset$)
10.0	3.3	95.0	3.9	-434.6	0.19	1.6E-02	2.5E-16
15.0	5.0	6.6	0.8	15.3	-0.36	1.1E-03	2.5E-16
20.0	6.6	1.1	0.9	15.9	-0.43	1.6E-03	2.5E-16
25.0	8.3	-2.9	0.2	16.4	0.40	1.9E-03	2.5E-16
30.0	10.0	-10.1	-0.3	17.2	0.01	2.9E-03	2.5E-16
35.0	11.6	-9.3	0.3	17.5	-0.26	5.4E-03	2.5E-16
40.0	13.3	-6.8	0.2	23.9	-0.48	2.8E-03	2.5E-16
μ						4.6E-03	2.5E-16

Table 5.4: SVD (order=1) offset fit statistics: fit from all data

End Data: SVD (order=1) raw maxima fit ($\{x,y,z\} \rightarrow \mu m$)							
M_{app} (g)	δz_{end}^{cant}	δx_{end}^{exp}	δy_{end}^{exp}	δz_{end}^{exp}	$2\pi\,\phi_{dev}$	χ_{fit}/ν	res($\emptyset$)
10.0	3.3	95.2	4.5	-433.8	0.09	7.3E-07	1.6E-14
15.0	5.0	24.3	1.1	-53.4	-0.21	2.3E-04	3.5E-16
20.0	6.6	-8.6	1.9	11.1	0.45	1.1E-03	1.1E-15
25.0	8.3	-13.3	0.4	8.1	0.30	9.3E-04	4.2E-16
30.0	10.0	-10.4	0.2	24.8	0.27	2.0E-03	1.2E-15
35.0	11.6	-12.7	-0.3	11.3	-0.02	2.7E-03	2.4E-16
40.0	13.3	-5.7	0.3	23.8	0.02	2.7E-03	4.9E-16
μ						1.4E-03	2.8E-15

End Data: SVD (order=1) raw phase unwind refit ($\{x,y,z\} \rightarrow \mu m$)							
M_{app} (g)	δz_{end}^{cant}	δx_{end}^{exp}	δy_{end}^{exp}	δz_{end}^{exp}	$2\pi\,\phi_{dev}$	χ_{fit}/ν	res($\emptyset$)
10.0	3.3	103.0	4.9	-432.3	-0.44	1.7E-02	5.2E-16
15.0	5.0	20.3	-0.2	-56.1	0.38	1.4E-02	5.2E-16
20.0	6.6	-6.4	1.2	11.0	0.20	8.0E-03	5.2E-16
25.0	8.3	-8.3	0.9	7.1	-0.44	7.6E-03	5.2E-16
30.0	10.0	-9.1	-0.1	18.2	-0.16	1.0E-02	5.2E-16
35.0	11.6	-12.8	0.0	4.3	0.46	8.5E-03	5.2E-16
40.0	13.3	-8.3	-0.1	22.5	0.29	8.5E-03	5.2E-16
μ						1.1E-02	5.2E-16

End Data: SVD (order=1) filtered phase unwind refit ($\{x,y,z\} \rightarrow \mu m$)							
M_{app} (g)	δz_{end}^{cant}	δx_{end}^{exp}	δy_{end}^{exp}	δz_{end}^{exp}	$2\pi\,\phi_{dev}$	χ_{fit}/ν	res($\emptyset$)
10.0	3.3	101.9	4.8	-429.1	0.48	1.3E-02	5.2E-16
15.0	5.0	16.6	-0.5	-57.8	-0.23	8.9E-03	5.2E-16
20.0	6.6	-3.7	0.6	13.0	0.11	1.5E-03	5.2E-16
25.0	8.3	-6.3	0.6	10.8	0.24	1.3E-03	5.2E-16
30.0	10.0	-9.9	1.0	26.4	0.00	3.4E-03	5.2E-16
35.0	11.6	-14.2	-0.5	6.4	-0.29	2.2E-03	5.2E-16
40.0	13.3	-8.9	0.6	26.0	-0.37	2.2E-03	5.2E-16
μ						4.7E-03	5.2E-16

Table 5.5: SVD (order=1) offset fit statistics: fit from end data

End Data: Constant offset raw maxima fit ($\{x, y, z\} \rightarrow \mu m$)							
M_{app} (g)	δz_{end}^{cant}	δx_{end}^{exp}	δy_{end}^{exp}	δz_{end}^{exp}	$2\pi\,\phi_{dev}$	χ_{fit}/ν	res($\emptyset$)
10.0	3.3	-1.4	-0.1	3.4	-0.02	2.1E-04	0.0E+00
15.0	5.0	-9.9	0.3	1.7	0.02	5.4E-04	0.0E+00
20.0	6.6	-13.4	1.3	2.6	-0.47	1.9E-03	0.0E+00
25.0	8.3	-16.6	1.1	3.6	0.29	1.9E-03	0.0E+00
30.0	10.0	-19.6	1.2	3.6	0.07	3.9E-03	0.0E+00
35.0	11.6	-22.4	0.7	4.4	0.38	4.5E-03	0.0E+00
40.0	13.3	-25.5	1.0	5.0	-0.38	4.8E-03	0.0E+00
μ						2.5E-03	0.0E+00

End Data: Constant offset raw phase unwind refit ($\{x, y, z\} \rightarrow \mu m$)							
M_{app} (g)	δz_{end}^{cant}	δx_{end}^{exp}	δy_{end}^{exp}	δz_{end}^{exp}	$2\pi\,\phi_{dev}$	χ_{fit}/ν	res($\emptyset$)
10.0	3.3	-2.0	0.3	3.2	-0.19	1.1E-02	0.0E+00
15.0	5.0	-10.3	0.6	1.6	-0.03	6.5E-03	0.0E+00
20.0	6.6	-13.5	1.2	2.6	-0.16	8.2E-03	0.0E+00
25.0	8.3	-16.8	1.1	3.5	-0.15	8.0E-03	0.0E+00
30.0	10.0	-19.8	1.2	3.6	-0.43	8.9E-03	0.0E+00
35.0	11.6	-22.6	0.8	4.4	-0.31	9.3E-03	0.0E+00
40.0	13.3	-25.5	0.9	5.1	-0.35	9.1E-03	0.0E+00
μ						8.8E-03	0.0E+00

End Data: Constant offset filtered phase unwind refit ($\{x, y, z\} \rightarrow \mu m$)							
M_{app} (g)	δz_{end}^{cant}	δx_{end}^{exp}	δy_{end}^{exp}	δz_{end}^{exp}	$2\pi\,\phi_{dev}$	χ_{fit}/ν	res($\emptyset$)
10.0	3.3	-2.1	-0.1	3.1	-0.01	5.4E-03	0.0E+00
15.0	5.0	-10.1	0.3	1.7	0.42	1.0E-03	0.0E+00
20.0	6.6	-13.1	1.1	2.7	-0.16	2.0E-03	0.0E+00
25.0	8.3	-16.7	1.1	3.6	0.45	2.0E-03	0.0E+00
30.0	10.0	-19.7	1.2	3.6	0.27	2.4E-03	0.0E+00
35.0	11.6	-22.5	0.8	4.4	-0.32	3.3E-03	0.0E+00
40.0	13.3	-25.4	1.1	5.1	0.37	3.4E-03	0.0E+00
μ						2.8E-03	0.0E+00

Table 5.6: Constant vector offset fit statistics: fit from end data

All Data: Cantilever raw maxima fit				
M_{app} (g)	M_{exp} (g)	M_{exp}/M_{app}	$2\pi\,\phi_{dev}$	χ_{fit}/ν
10.0	11.08	1.108	-0.489	0.00027
15.0	15.05	1.003	+0.093	0.00576
20.0	22.16	1.108	+0.245	0.00694
25.0	28.89	1.156	+0.102	0.00634
30.0	31.17	1.039	-0.230	0.03867
35.0	38.37	1.096	+0.312	0.00671
40.0	42.29	1.057	+0.347	0.02221
μ		1.081		0.01242
σ		0.051		

All Data: Cantilever raw phase unwind refit				
M_{app} (g)	M_{exp} (g)	M_{exp}/M_{app}	$2\pi\,\phi_{dev}$	χ_{fit}/ν
10.0	11.19	1.119	-0.491	0.00539
15.0	14.97	0.998	+0.081	0.00647
20.0	22.03	1.102	+0.239	0.00678
25.0	28.66	1.147	+0.106	0.00896
30.0	30.86	1.029	-0.238	0.01042
35.0	37.88	1.082	+0.329	0.01208
40.0	42.53	1.063	+0.315	0.01352
μ		1.077		0.00909
σ		0.052		

All Data: Cantilever filtered phase unwind refit				
M_{app} (g)	M_{exp} (g)	M_{exp}/M_{app}	$2\pi\,\phi_{dev}$	χ_{fit}/ν
10.0	11.18	1.118	-0.488	0.00106
15.0	14.91	0.994	+0.074	0.00350
20.0	21.83	1.091	+0.233	0.00307
25.0	28.64	1.146	+0.094	0.00239
30.0	30.82	1.027	-0.252	0.00727
35.0	37.90	1.083	+0.321	0.00662
40.0	42.42	1.061	+0.309	0.01085
μ		1.074		0.00497
σ		0.052		

Table 5.7: Cantilever fit statistics: fit from all data

End Data: Cantilever raw maxima fit				
M_{app} (g)	M_{exp} (g)	M_{exp}/M_{app}	$2\pi\,\phi_{dev}$	χ_{fit}/ν
10.0	11.08	1.108	-0.489	0.00027
15.0	15.93	1.062	+0.003	0.00103
20.0	22.82	1.141	+0.166	0.00320
25.0	29.28	1.171	+0.058	0.00192
30.0	33.02	1.101	-0.445	0.00451
35.0	38.74	1.107	+0.270	0.00493
40.0	43.42	1.085	+0.228	0.00659
μ		1.111		0.00321
σ		0.036		

End Data: Cantilever raw phase unwind refit				
M_{app} (g)	M_{exp} (g)	M_{exp}/M_{app}	$2\pi\,\phi_{dev}$	χ_{fit}/ν
10.0	11.18	1.118	-0.489	0.00676
15.0	15.99	1.066	-0.002	0.00709
20.0	22.71	1.136	+0.165	0.00836
25.0	29.16	1.166	+0.057	0.00835
30.0	32.70	1.090	-0.425	0.01113
35.0	38.58	1.102	+0.272	0.01120
40.0	43.00	1.075	+0.241	0.01125
μ		1.108		0.00916
σ		0.035		

End Data: Cantilever filtered phase unwind refit				
M_{app} (g)	M_{exp} (g)	M_{exp}/M_{app}	$2\pi\,\phi_{dev}$	χ_{fit}/ν
10.0	11.26	1.126	-0.494	0.00117
15.0	15.92	1.061	-0.002	0.00147
20.0	22.59	1.129	+0.170	0.00241
25.0	29.12	1.165	+0.053	0.00179
30.0	32.42	1.081	-0.412	0.00551
35.0	38.44	1.098	+0.282	0.00685
40.0	42.88	1.072	+0.245	0.00604
μ		1.105		0.00361
σ		0.037		

Table 5.8: Cantilever fit statistics: fit from end data

free end causes the deformation, applied mass gives the sensible free parameter. All fits returned expected deformation measures to better than 4% standard error, with fits on end data only doing slightly better than those utilizing data from more of the cantilever surface.

Theoretical interferograms for the cantilever fits match experimental results better near the free end than toward the constrained end of the cantilever. Such characteristics are sensible, in that experimental boundary conditions at the fixed end failed to match those of the ideal cantilever. One should not have expected that theoretical boundary conditions could be perfectly arranged under experimental constraints. One expects intuitively, however, that induced errors should lose effect, as measurements become increasingly remote from imperfectly modelled boundaries.

The cantilever's constrained end condition was complicated by several factors. The beam was attached to a ledge via a brass bolt $\sim 9\ mm$ back from the limit. Moreover, a $3.175\ mm$ hole punctured the $2.54\ mm \times\ 155.5\ mm$ surface, centred $15.5\ mm$ back from the ledge. While bolt location and the hole were accounted for in the theoretical model, no allowance was made for strain at the bolt. Comparison of theoretical and experimental interferogram images shows, that the rather complicated experimental boundary condition failed to match perfectly, with that of the corrected theoretical model.

One finds in ironic contrast, as illustrated by figure 5.29, that interferograms simulated from *SVD* fits over the full cantilever match better with the experimental images than do the respective cantilever simulations. The

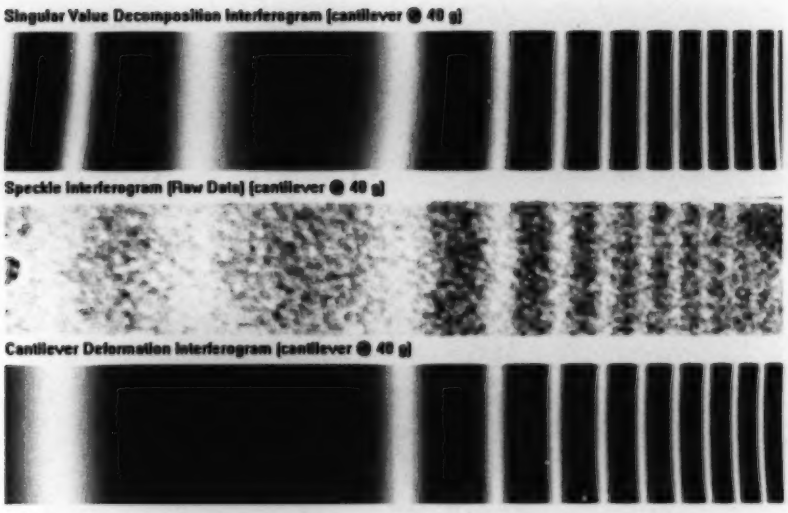

Singular Value Decomposition Interferogram (cantilever @ 40 g)
Speckle Interferogram (Raw Data) (cantilever @ 40 g)
Cantilever Deformation Interferogram (cantilever @ 40 g)

Figure 5.29: Fit interferogram comparison: *SVD* vs. Cantilever

seeming paradox posed by such a result does however yield to analysis. Quite simply, the *SVD* results do fit the experimental data better than do the cantilever fits. By virtue of variation in altogether far too many free parameters, the first order *SVD* algorithm found deformations, which were at once theoretically optimal and experimentally nonsensical. The comparatively suboptimal theoretical cantilever fits produced realistic results, primarily because they were so constrained, to model cantilever deformations.

One may verify in the results, of data tables 5.4 through 5.8, that *SVD* fits generally outperformed their respective cantilever counterparts. Comparison of tabulated values for χ/ν (variance per degree of freedom) shows this. For example, raw data fit *SVD* results show average variances $\chi/\nu = 1.4 \times 10^{-3}$,

both over the entire cantilever, and on fits near the end alone. Cantilever deformation fits on the same data show average variances of 1.2×10^{-2} over the full surface, and 3.2×10^{-2} near the free end, away from the imperfect boundary condition.

One concludes that the *SVD* method gives unreliable results, when fits are subject to real experimental noise levels, present in the data of this study. One might well ask whether data with lower noise levels would have yielded to more than a single parameter fit. As demonstrated in the simulations of subsections 5.1.3 and 5.1.4, noise levels at single precision floating point accuracy (5 significant figures) do allow for *SVD* type fits up to even third order. The obvious question, is whether experimental noise levels can be significantly reduced, to a degree where deformations of more than lowest order may be fit.

The effect of statistical noise and its limitation has been examined in detail, as outlined in subsection 2.1.3. The limits to statistical observation should be examined from the same theoretical vantage. The hypothesis of optimal data weighting, (restricted to fringe peaks alone,) as expressed in equation 2.20, may be examined directly. One may attempt data refits, on equation 2.26 type phase unwound raw or filtered data. One may also derive the effect of global thresholding, on phase variance of local maxima, thereby inferring a theoretical limit to speckle-interferometric phase identification.

The prediction that optimal fits derive from data restriction to fringe maxima, per equation 2.20, was tested against each of the respective data fits.

As expected, all attempts to refit phase unwound raw data resulted in greatly increased values for χ/ν (variance per degree of freedom). Attempts to refit phase unwound filtered data failed to improve results on a similar measure.

Given the extreme nature of the functional image fit, outlined in equation 2.24, which has been applied in generating the respective phase unwinding terms, the effects of over-tuned data should be observable. One should expect unreasonably good refits, even with equal global data weighting. Consistency of interferogram data, with the primary fit, has after all been imposed quite artificially, in manufacturing the new data. Since no such dramatic improvement follows, one must conclude that results concur better with the hypothesis of equation 2.20. Optimal results do seem to come from weighted fits, restricting data solely to interference maxima.

The hypothesis of the linearized cross-correlation error, expressed in equation 2.9, which led to the assignment of optimal data weights, per equation 2.20, does in fact yield a final prediction. One may derive the variance associated with identification of maxima, as a function of the global threshold used for the purpose of finding these. Whether a Gaussian or rectangular distribution is applied, with the theoretical linearized variance of equation 2.9, one achieves roughly the same result. Identification of interference fringe peaks shows an associated theoretical variance as a function of test threshold.

The theoretical distribution may be readily tabulated, and appears displayed for 5×5 cross-correlation sample regions, in figure 5.30. Identification of fringe maxima by threshold is subject to a sharply defined "sweet spot".

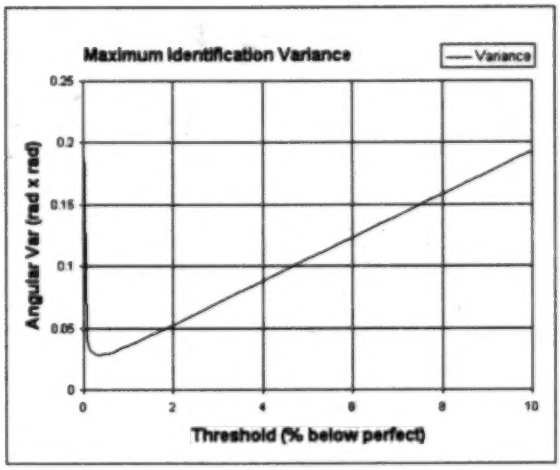


Figure 5.30: Theoretical peak location variance as function of threshold

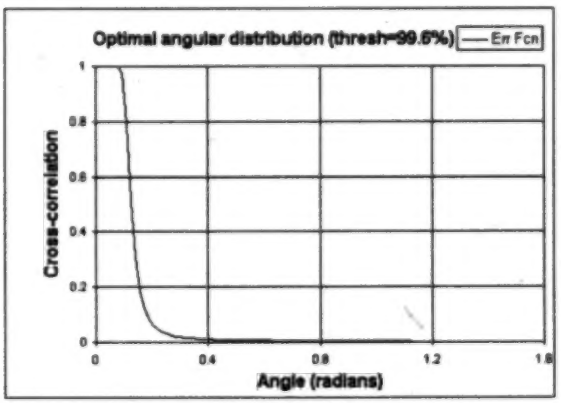


Figure 5.31: Theoretical peak distribution with optimal thresholding

Optimal thresholding occurs at 99.6% correlation, with $\sim$ 253/256 8-bit imaging equivalent.

The ideal distribution for the experiment at hand appears in figure 5.31. The optimal variance in this case was found to be $\sigma^2_{\phi\phi} = 0.0285$. Due to systematic errors of the types outlined in equation 2.23, however, one may in practice only approach the ideal value. A more realistic working estimate for the expected variance corresponds to the limit value ($\tau \to 1$) of unit threshold. Phase identification variance approaches $\sigma^2_{\phi\phi} = 0.213$ in this limit.

One may derive an asymptotic equivalent for the limiting result, by assuming small angle approximations on the cosines, and high z-score asymptotic error functions. With nominal $\kappa = 1/2$ in equation 1.5 one finds a simple steepest descent (constant phase) approximation:

$$\sigma^2_{\phi\phi} \quad \xrightarrow{\tau \to 1} \quad \sim \frac{8}{8+N} \tag{5.2}$$

Experimental value $N = 5 \times 5$ gives $\sigma^2_{\phi\phi} \approx 0.2424$, in good agreement with the detailed calculation per figure 5.30.

Results for theoretical phase identification variance may be compared with experimental values for χ/ν (variance per degree of freedom). The former results as angular radian variances, and the latter wave number variances, may both be converted to standard deviations in degrees. Optimal and working theoretical errors correspond respectively to 9.7° and 28.2°. A nominal working (wave number) variance of 1.5×10^{-3} has equivalent standard error $\sigma = 13.9°$. Experimental values fall within the theoretical range.

Analysis of data variances indicates that observed experimental errors closely approach the limits of measurement for a speckle interferometer. Interference phase measurement precision cannot improve, by the several orders of magnitude required, for successful execution of a multi-parameter interferogram deformation fit. As a result, the general inverse problem of this thesis remains beyond the capabilities of a simple interferometer. General deformation measurment requires more complex interferometric techniques than those of the current research.

Simple deformations however, with only a few free parameters, may be measured with high precision. Such simplicity derives by constraining inverse problems to real physical models. Model based fitting procedures, as carried out here for the cantilever deformation, were originally demonstrated, with practical and theoretical refinements, by *Schmitt and Hunt* in the late 1990s.[SCHM 97][SCHM 98][SCHM 00] Practical application there involved hole-drilling residual stress determinations. Theoretical refinements in the last paper included optimization search for best integral phase angle offset.

Generalizations made in the current research show that many similar applications have promise. Wherever simply parametrized smooth physical models apply to microscopic deformations, similar fits might be attempted; solution for deformation parameters, including phase angle offset, always reduces to simple algebra. Many simple parametrized deformations should yield to high precision interferometric measurment. Experimental investigation might reveal which deformation types yield meaningful results.

5.3.1 Raw (5 × 5 sample) cross-correlations

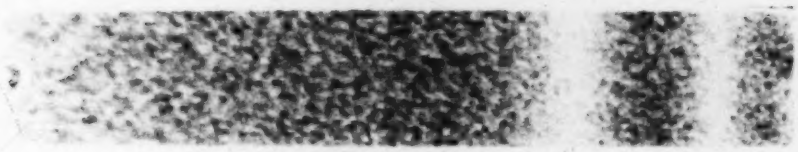


Figure 5.32: Cantilever deformation under 10g

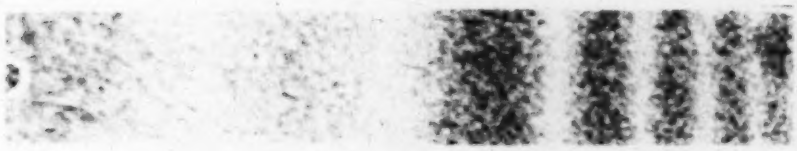


Figure 5.33: Cantilever deformation under 20g

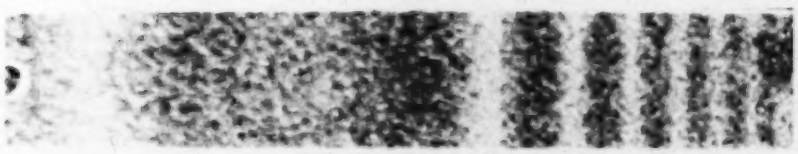


Figure 5.34: Cantilever deformation under 30g

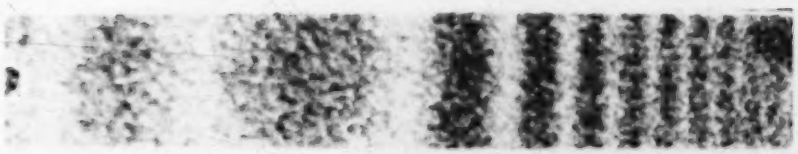


Figure 5.35: Cantilever deformation under 40g

5.3.2 Filtered cross-correlations

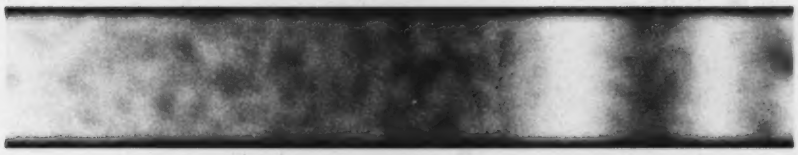


Figure 5.36: Cantilever deformation under 10g

Figure 5.37: Cantilever deformation under 20g

Figure 5.38: Cantilever deformation under 30g

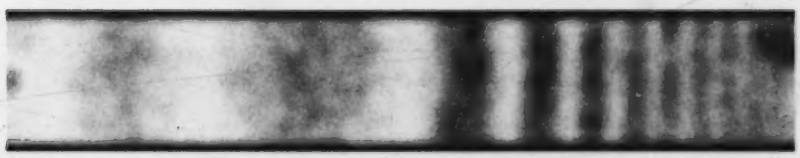


Figure 5.39: Cantilever deformation under 40g

5.3.3 Theoretical deformations

Figure 5.40: Cantilever deformation under 10g

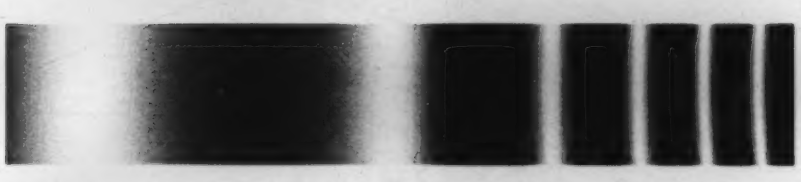


Figure 5.41: Cantilever deformation under 20g

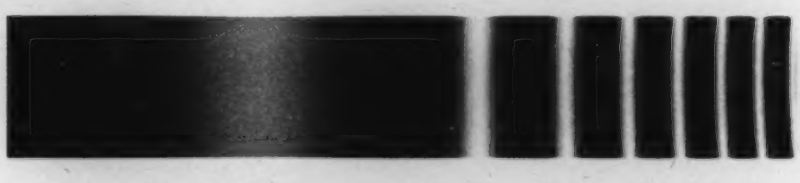


Figure 5.42: Cantilever deformation under 30g

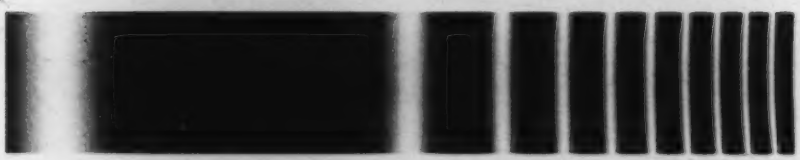


Figure 5.43: Cantilever deformation under 40g

5.3.4 Raw interferograms (scaled & offset)

Figure 5.44: Cantilever deformation under 10g

Figure 5.45: Cantilever deformation under 20g

Figure 5.46: Cantilever deformation under 30g

Figure 5.47: Cantilever deformation under 40g

5.3.5 Filtered interferograms (scaled & offset)

Figure 5.48: Cantilever deformation under 10g

Figure 5.49: Cantilever deformation under 20g

Figure 5.50: Cantilever deformation under 30g

Figure 5.51: Cantilever deformation under 40g

Part III

Conclusions

Chapter 6

What has been shown?

Electronic speckle phase interferometry *(ESPI)* has been investigated as a stochastic phenomenon and demonstrated as a tool in analysis of surface deformations. Some benefits and drawbacks of *ESPI* interferometric deformation observation were revealed in this research.

Classical stochastic methods found natural application in the field of speckle interferometry. Expected intensity distributions for possibly partially polarized laser scatter were evaluated. Interference effects between images were also subjected to stochastic analysis. Results from such analysis showed substantial agreement with the simple model of *Jones and Wykes.*[JONE 83] Adjustments, due to interactions involving polarizations and surface incidence, left the perception of interference maxima unaltered, with minor observable effects predicted elsewhere. Speckle interferograms were found, in theory and experiment, to show resilience with respect to variation of pixel magnification, and even under loss of optical focus.

Theory predicted a purely statistical error in cross-correlation sampling. This error literally defined the scope of possibility for experimental speckle interferometry. Variance analysis showed that constructive interference maxima hold all the information content in an interferogram, whereas other suboptimal interferogram data have relatively negligible data weight contributions, in optimal fitting procedures. Experimental tests consistently agreed with theoretical application of this linearized error.

An attempt was made to automate deformation analysis. Data filtering and identification of interferogram maxima reduced to a purely mechanical computation. Theory indicated, that application of a rectangular median filter, followed by simple global thresholding, sufficed for identification of the local maxima, comprising all relevant (optimal) interferogram data. Some aspects of the analysis were found to be more difficult to automate. For example, identification and assignment of phase angles at local maxima remains a nontrivial task.

A general method for deformation analysis, depending solely on single simple *ESPI* interferograms was presented. Relying on singular value decomposition methods, to deal with symmetries inherent to the experimental geometry, the method provided a mechanism for identifying and rationalizing interferometrically invisible displacement fields. Tests on artificial and relatively noise-free phase data revealed, that such a method was workable, provided sufficiently clean image interferograms.

This same method could be extended, to allow for tuning of experimental

geometry. Utilizing a single interferogram, for a rigid (Killing vector) displacement, locations of interferometer scattering sources could be iteratively corrected from approximate experimental measurements. The algorithm prescribes a simple mechanism for elimination of systematic errors from interferometer deformation measurements.

Limitations of multi-parameter deformation fits were explored and contrasted with simple single parameter physical models. This comparison involved examination of a simple cantilever experiment. Multi-parameter fits were found to fail by returning nonsensical deformation fields; The simple single parameter cantilever equation meanwhile gave consistently reasonable results. When deformations were modelled against original interferograms, however, roles reversed: the nonsensical multi-parameter deformations gave better image fits than did deformations modelled for the ideal cantilever.

Such failure was in fact expected, due partially to the limitations imposed by cantilever geometry, but due primarily to natural precision limitations in interferometer phase measurement. The natural limitation to precision derives from the aforementioned purely statistical error in cross-correlation sampling. Phase angle identification error was modelled from this phenomenon, giving results consistent with experimental values.

An obvious conclusion follows: observed experimental errors closely approach the limits of measurement for a simple speckle interferometer; measurement precision, therefore, cannot improve, by the several orders of magnitude required, for success in multi-parameter deformation fits. The

general problem of interferometric deformation analysis, exlored in this research, clearly requires more sophisticated techniques than the simple *ESPI* method of this thesis.

Simple deformation analysis methods based on physical models do however show promise. By variation of only a few parameters, such models limit the effects of statistical noise on results. High precision measurements were shown to be possible in the simple case of the cantilever deformation. Present results also show good consistency with prior simple model-based analysis.[SCHM 97][SCHM 98][SCHM 00] Since many common deformations may be simply parametrized, similar methods should be applicable, to high precision interferometric measurment, in a diversity of practical applications.

Bibliography

[BHAT 86] A.B. Bhatia & R.N. Singh, *Mechanics of deformable media,* (Bristol: Adam Hilger, 1986).

[CHAN 89] S. Chandrasekhar, *Stochastic, statistical, and hydromagnetic problems in physics and astronomy,* (Chicago: University of Chicago Press, 1989).

[CHUN 83] K.L. Chung & R.J. Williams, *Introduction to stochastic integration,* (Boston: Birkhäuser, 1983).

[DAIN 89] J.C. Dainty, ed., *Laser Speckle and related phenomena,* 2d ed. (Berlin: Springer-Verlag, 1989).

[EDWA 85] C.H. Edwards Jr. & D.E. Penney, *Elementary differential equations with applications,* 3d ed. (Englewood Cliffs, New Jersey: Prentice-Hall, 1985).

[EMER 89] M. Emery & P.A. Meyer, *Stochastic calculus in manifolds,* (Berlin: Springer-Verlag, 1989).

[ENNO 89] A.E. Ennos, *Laser interferometry* in J.C. Dainty, ed., *Laser speckle and related phenomena,* 2d ed. (Berlin: Springer-Verlag, 1989) 207-211.

[FELD 96] R.M. Feldman & C. Valdez-Flores, *Applied probability & stochastic processes,* (Boston: PWS Pub. Company, 1996).

[FELL 68] W. Feller, *An introduction to probability theory and its applications,* 3d ed. (New York: Wiley, 1968).

[FREU 62] J.E. Freund & R.E. Walpole, *Mathematical statistics,* 3d ed. (Englewood Cliffs, New Jersey: Prentice-Hall, 1962).

[GLIM 81] J. Glimm & A. Jaffe, *Quantum physics, a functional integral point of view,* 2d ed. (New York: Springer-Verlag, 1981).

[GOOD 89] J.W. Goodman, *Statistical properties of laser speckle* in J.C. Dainty, ed., *Laser speckle and related phenomena,* 2d ed. (Berlin: Springer-Verlag, 1989) 21-26.

[HECH 74] E. Hecht & A. Zajac, *Optics,* (Reading, Massachusetts: Addison-Wesley, 1974).

[ITÔ 74] K. Itô & H.P. McKean, *Diffusion processes and their sample paths,* (Berlin: Springer-Verlag, 1974).

[JACK 62] J.D. Jackson, *Classical electrodynamics,* 2d ed. (New York: Wiley, 1962).

[JONE 83] R. Jones & C. Wykes, *Holographic and speckle interferometry,* (Cambridge: Cambridge University Press, 1983).

[KERN 88] B.W. Kernighan & D.M. Ritchie, *The C programming language,* 2d ed. (Englewood Cliffs, New Jersey: Prentice-Hall, 1988).

[LEEN 70] J.A. Leendertz, *Interferometric displacement measurement on scattering surfaces utilizing the speckle effect* (1970) 3 J. Phys. E Sci. Instrum. 214-218.

[MATH 70] J. Mathews & R.L. Walker, *Mathematical methods of physics,* 2d ed. (Reading, Massachusetts: Addison-Wesley, 1970).

[PATH 72] R.K. Pathria, *Statistical mechanics,* (Oxford: Pergamon Press, 1972).

[PRES 92] W.H. Press et al., *Numerical recipes in C: the art of scientific computing,* 2d ed. (Cambridge: Cambridge University Press, 1992).

[SCHM 89] D.R. Schmitt et al., *In-situ holographic elastic moduli measurements from boreholes* (1989) 54 Geophysics 468-477.

[SCHM 97] D.R. Schmitt & R.W. Hunt, *Optimization of fringe pattern calculation with direct correlations in speckle interferometry* (1997) 36 Applied Optics 8848-8857.

[SCHM 98] D.R. Schmitt & R.W. Hunt, *Model-based invertion of speckle interferometer fringe patterns* (1998) 37 Applied Optics 2573-2578.

[SCHM 00] D.R. Schmitt & R.W. Hunt, *Inversion of speckle interferometer fringes for hole-drilling residual stress determinations* (2000) 40 Experimental Mechanics 129-137.

[SCHU 80] B.F. Schutz, *Geometrical methods of mathematical physics,* (Cambridge: Cambridge University Press, 1980).

[SEN 93] P.K. Sen & J.M. Singer, *Large sample statistics: an introduction with applications,* (New York: Chapman & Hall, 1993).

[STRA 63] R.L. Stratonovich, *Topics in the theory of random noise,* (New York: Gordon and Breach, 1963).

[SVOZ 93] K. Svozil, *Randomness & undecidability in physics,* (Singapore: World Scientific, 1993).

[TAYL 98] H.M. Taylor & S. Karlin, *An introduction to stochastic modelling,* 3d ed. (San Diego: Academic Press, 1998).

Appendices

Appendix A

Probabilities

A.1 Distributions and densities

The principal objective of the study of probability is precise prediction of the nature and scope of outcomes in sets of random experiments.[FREU 62] Statistics is in this sense a purely empirical field. Detailed analyis of random events from such a perspective does however require a strict theoretical foundation. The fundamental connection between empirical and theoretical statistics derives in the concept of the *distribution*.

The theoretically assumed *global distribution* includes all possible outcomes. A *probability* is assigned to each outcome, proportional to its frequency of occurence in *independent trials*. Experimental distributions derive conversely from data measurements of *empirical samples*. Empirical samples however rarely describe an entire global *sample space*. Rigorous scrutiny is thus required in extrapolation of results from incomplete samples to respec-

tive complete global populations.

Exhaustive research exists investigating the relationship between *sample* and *global* distributions. The *law of large numbers* suggests empirical sample distributions should in some sense approach global distribution values as independent trial iterations increase without bound. At the other extreme, study of small samples introduces a host of new and diverse complexities. Related issues have real bearing on research in this thesis, as illustrated by the effect of pure sampling error in cross-correlations. Investigation of this statistical error (see subsection 2.1.3) reveals drastic limitations on information content (see subsection 2.1.5) of *ESPI* interferograms.

Rigorous introduction to statistical probability requires development of further concepts, including the relationship between the *partition* and the *random variable*. Precise definition of a global distribution requires that random outcomes be clearly distinguishable in some natural view: each experimental outcome should correspond to a single unique theoretical outcome. Such identification defines the concept of the *partition*. If a partition is parametrized or indexed, values of parameters or indices can thus be measured as *random variables* in experimental trials. One may generalize the concept of random variables to include all functions of outcomes in random experiments on a partition.

Strict theoretical development of statistical probability provides a solid foundation for practical measurement in empirical experiments. Illuminating as such study might be, detailed study of the theory of probability is beyond

the scope of this work. Some results can be quoted, however, as definitions respected throughout the thesis at hand.

The principal definition is that of the *probability distribution*. General partitions may be parametrized in some combination of discrete and continuous sets. Classical literature distinguishes between the purely discrete and purely continuous cases by defining respectively the *distribution* (discrete case) and *density* (continuous case).[FREU 62] Modern treatments based in measure theory view such distinction as artificial.[FELD 96][TAYL 98] The general term *distribution* is thus applied to both discrete and continuous cases throughout this thesis.

In any case, probability derives as a non-negative function on a partition. Probability functions of possible experimental outcomes x necessarily satisfy completeness criteria by the respective normalization conditions:

$$\sum_x P_x = 1 \quad \text{or} \tag{A.1}$$

$$\int_x dx\ P(x) = 1 \tag{A.2}$$

The (discrete) sum in equation A.1 and the (continuous) integral in equation A.2 cover fully the respective partitions of possible independent outcomes x. One further defines the concept of *mathematical expectation* by the brackets $\langle \ldots \rangle$ for general functions of random variable x as follows:

$$\langle f_x \rangle_x = \sum_x P_x \cdot f_x \quad \text{or} \tag{A.3}$$

$$\langle f(x) \rangle_x = \int_x dx\ P(x) \cdot f(x) \tag{A.4}$$

One may write for numerical random variable x some immediate definitions. The *mean* μ and *variance* $\sigma_{xx}^2 = var(x)$ are as follows:

$$\mu = \quad \langle x \rangle_x \quad \text{and} \tag{A.5}$$

$$\sigma_{xx}^2 = \langle (x-\mu)^2 \rangle_x = \langle x^2 \rangle_x - \langle x \rangle_x^2 \tag{A.6}$$

The *median* gives an intuitive midpoint for x in a distribution or density:

$$\tilde{\mu} = \min\{z : \langle \theta(z-x) \rangle_x \geq \frac{1}{2}\} \quad \text{where} \quad \theta(\mathrm{t}) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases} \tag{A.7}$$

The *mode* corresponds to the set of most frequent occurrences for a distribution or density:

$$\bar{\mu} = \{z : \forall x \begin{cases} P_x \leq P_z & \text{discrete} \\ P(x) \leq P(z) & \text{continuous} \end{cases}\} \tag{A.8}$$

One should note that while the mean, median and support of the mode might in particular cases coincide, no such agreement is generally required.

One may moreover measure a distribution in terms of its defining expectation values by considering an experimental *empirical sample distribution.* The sample corresponds to a proper subset of all possible random experiments. One considers the measured expectations or empirical means on the sample to be *estimators* of their global counterparts on the total set of possible experiments.

Expectations measured on samples are also random variables. One may study the manner in which these in turn are distributed. Following analysis of this sort, one describes an estimator as *unbiased* when its expectation

matches the related global quantity. For example, the unbiased estimators from a sample of size N for global mean and variance, on any distribution or density, are given respectively as follows:

$$\mu^S = \frac{1}{N}\sum_{s\in S} x_s \tag{A.9}$$

$$\sigma_{xx}^{2\,S} = \frac{1}{N-1}\sum_{s\in S}(x_s - \mu^S)^2 \tag{A.10}$$

One may further conduct comparative studies amongst results from sampling experiments. Extrapolations involving *hypothesis testing* are considered in the chapter B appendices.

In conclusion of this cursory introduction to probabilty, one may tabulate a few common density functions. Distributions utilized throughout this thesis include the following:[FREU 62]

$$\kappa(x;a,b) = \begin{cases} \frac{1}{|a-b|} & x \in [a,b] \\ 0 & \text{elsewhere} \end{cases} \quad \text{constant or rectangular} \tag{A.11}$$

$$g(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{Gaussian or normal} \tag{A.12}$$

$$\gamma(x;\alpha,\beta) = \frac{1}{\beta^\alpha\,\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\frac{x}{\beta}} \quad \text{gamma density} \tag{A.13}$$

$$\chi(x;\nu) = \gamma(x;\, \alpha = \tfrac{\nu}{2},\, \beta = 2) \quad \text{chi-square density} \tag{A.14}$$

$$c(\phi;\Phi) = \tfrac{\Phi}{\pi}\,(\Phi^2 + \phi^2)^{-1} \quad \text{Cauchy density} \tag{A.15}$$

A.2 Means and medians

At the conclusion of a test, the experimentalist often questions, whether data might have been better analyzed, to extract more precise results. A couple of results are worth presenting in this context. First, even with perfect unbiased data, one would approach a limit of diminishing returns. Second, theoretical optimality of estimation figures as a nontrivial concept, when confronted with real corrupt and noisy data.

Given a variable one wishes to measure, one might reasonably wonder how to treat a set of measurements each of which returns an estimate of the quantity. One may show, by virtue of the *Cramér-Rao* inequality, that for asymptotically independent samples from Gaussian distributions, the usual sample mean $\langle x \rangle$ gives the best available (minimum variance) estimate of the true population mean.

The measurement of an experimental sample mean may be somewhat finessed in precision. One attempts a weighted mean, and finds, by simple projection or variational analysis, the optimal measurement weights. The optimal mean measurement, thus calculated, has the minimum variance for any choice of data weights:

$$\begin{aligned} \bar{x} &= \sum w_n x_n \quad \text{with constraint } \sum w_n = 1 \text{ gives} \\ \sigma_{xx}^2 &= \sum w_n \cdot (x_n - \bar{x})^2 \quad \text{with optimal weight solution} \\ w_n &= \frac{1}{\sigma_n^2} \cdot (\sum_k \frac{1}{\sigma_k^2})^{-1} \quad \text{where } \sigma_n^2 = \langle (x_n - \bar{x})^2 \rangle \end{aligned} \tag{A.16}$$

One may derive maximum likelihood variance for sample mean $\bar{x}$ as follows:

$$\sigma^2_{\bar{x}\bar{x}} = (\sum_k \frac{1}{\sigma_k^2})^{-1} \tag{A.17}$$

One infers that, unless individual measurement errors show disparate scales of magnitude, such optimization will not yield significant improvement from the corresponding unweighted result.

When data distributions are punctuated with spike noise error signals, a *mixture* distribution results. In such mixtures, the result of any mean calculation may be compromised. Consider for example, the case of a Gaussian distribution with mean μ_0 and standard deviation σ_0, corrupted by Gaussian spike noise of mean μ_{err} and standard deviation σ_{err} comprising a fractional measure ϵ of the combined global population:

$$p(x) = \epsilon \cdot g(x; \mu_{\text{err}}, \sigma_{\text{err}}) + (1-\epsilon) \cdot g(x; \mu_0, \sigma_0) \tag{A.18}$$

The mean and standard deviation, calculated for the corrupted net signal, are given respectively by:

$$\begin{aligned} \mu &= \epsilon \cdot \mu_{\text{err}} + (1-\epsilon) \cdot \mu_0 \\ \sigma^2 &= \epsilon \cdot \sigma^2_{\text{err}} + (1-\epsilon) \cdot \sigma_0^2 + \epsilon(1-\epsilon) \cdot (\mu_{\text{err}} - \mu_0)^2 \end{aligned} \tag{A.19}$$

Deviations from the requisite values of μ_0 and σ_0 are potentially boundless. Note further, that the results of equation A.19 may be generalized from simple Gaussians, to all distributions for which the requisite moments exist. One clearly has to find an experimental *estimator* which shows more resilience under experimental spike noise mixture.

Much statistical study has been devoted to investigation of the problem of *consistent* and *sufficient* estimators. In theory, one may construct high efficiency unbiased estimators which work well with asymptotically large samples. Practically, dramatic improvement may be demonstrated, even restricted to small samples, with a relatively simple solution.

The median of a distribution (defined in equation A.7) permits simple analysis. If continuous random variable x has density $p(x)$, then the median for a sample of odd size $2N+1$ distributes as density $m_N(x)$ (through simple multinomial combinatorics) by the equation:

$$m_N(\tilde{x}) \;=\; \frac{(2N+1)!}{N!\,N!}\,\{\int_{-\infty}^{\tilde{x}} dx\,p(x)\,\}^N \cdot p(\tilde{x}) \cdot \{\int_{\tilde{x}}^{+\infty} dx\,p(x)\,\}^N \qquad \text{(A.20)}$$

One may eliminate the noncompact integrations, with aid of population median $\tilde{\mu}$, and intuit an asymptotic behaviour as follows:

$$\begin{aligned} m_N(\tilde{x}) \;&=\; \tfrac{(2N+1)!}{N!\,N!}\,\{\tfrac{1}{2} + \textstyle\int_{\tilde{\mu}}^{\tilde{x}} dx\,p(x)\}^N \cdot p(\tilde{x}) \cdot \{\tfrac{1}{2} - \textstyle\int_{\tilde{\mu}}^{\tilde{x}} dx\,p(x)\}^N \qquad \text{(A.21)} \\ &=\; \tfrac{(2N+1)!}{2^{2N}\,N!\,N!}\,p(\tilde{x})\,\{1 - [2\cdot\textstyle\int_{\tilde{\mu}}^{\tilde{x}} dx\,p(x)]^2\}^N \\ &\sim\; \tfrac{(2N+1)!}{2^{2N}\,N!\,N!}\,p(\tilde{x})\,\exp(\,-\tfrac{1}{2}\,[2\sqrt{2N}\textstyle\int_{\tilde{\mu}}^{\tilde{x}} dx\,p(x)]^2\,) \\ &\sim\; 2\sqrt{\tfrac{N}{\pi}}\,p(\tilde{\mu})\,\exp(\,-\tfrac{1}{2}\,[2\sqrt{2N}\,p(\tilde{\mu})\,(\tilde{x}-\tilde{\mu})]^2\,) \\ &\sim\; g(\,\tilde{x};\,\mu=\tilde{\mu},\,\sigma_N=\tfrac{1}{2\sqrt{2N}\,p(\tilde{\mu})}\,) \end{aligned}$$

The above derivation does not withstand rigorous scrutiny. Precise treatment of the limits does however affirm the simple result. Motivated by equation A.21, one defines random variable $z = \dfrac{x-\tilde{\mu}}{\sigma_N}$, with $\sigma_N = \dfrac{1}{2\sqrt{2N}\,p(\tilde{\mu})}$. New random variable z distributes precisely as $M_N(z) \;=\; \sigma_N\,p(\tilde{\mu}+\sigma_N z)$,

with following exact and limiting expressions:

$$M_N(z) = \frac{(2N+1)!}{2^{2N+1}\sqrt{2N}\,N!^2}\,\frac{p(\tilde{\mu}+\sigma_N z)}{p(\tilde{\mu})}\,\{1-\frac{1}{2N}[\int_0^z dz\,\frac{p(\tilde{\mu}+\sigma_N z)}{p(\tilde{\mu})}]^2\}^N \qquad \text{(A.22)}$$
$$\sim\ g(\,z;\,\mu=0,\sigma=1\,) = \frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}z^2}$$

The standard normal limit in equation A.22 holds provided z is held bounded and with $p(x)$ nonvanishing and continuous in a neighbourhood about median value $x = \tilde{\mu}$.

Distribution of the sample median is thus asymptotically Gaussian. One may rank the *asymptotic efficiency* of the median with respect to the mean on samples from a Gaussian population:

$$E = \lim_{N->\infty} \frac{var(\tilde{x})_{2N+1}}{var(\bar{x})_{2N+1}} = \frac{\pi}{2} \qquad \text{(A.23)}$$

One infers that the sample mean gives a somewhat more reliable estimate of the population mean/median than does the respective sample median.

When Gaussian data are corrupted by mixture with spike noise, (as per equation A.18,) however, the roles actually reverse. Possible deviation of the experimental sample mean from the requisite global quantity remains inestimable. The median however cannot vary by more than a fixed measure:

$$\frac{|\tilde{\mu}-\mu_0|}{\sigma_0} \leq \sqrt{2}\,\mathrm{erf}^{-1}(\frac{2\epsilon}{1-\epsilon}) \quad \text{where} \qquad \text{(A.24)}$$
$$\mathrm{erf}\,(s) = \frac{2}{\sqrt{\pi}}\int_0^s dt\,e^{-t^2}$$

Strictly considered, the deviation does derive from a systematic bias. One finds though, in cases involving small samples, that such a bias often appears insignificant when compared with Gaussian signal noise. Consistent

identification of locally dominant populations is thus quite achievable, via the sample median, provided one accepts the systematic estimate bias.

The above analysis may be extended to still more extreme limits. For some pathological distributions, such as the Cauchy density of equations 2.22 and A.15, moments such as the mean μ and variance σ^2 actually fail to exist. One finds, via equation A.17, that the sample mean fails to target the Cauchy central peak. The sample median does however find its mark with that distribution, as one infers from equations A.21 and A.22.

One finds that real analysis on experimental data often conceals subtleties of the types discussed in the foregoing. Some insight into the statistical nature of a particular measured observable often allows for better measurement of related quantities. In practice, however, one often discovers problems by exactly the reverse process: one establishes the limits of precision through measurements that "should have worked better", and in the mandatory forensics following such "failed" experiments. Fortunately one finds, forensic investigations into limitations of experimental precision often offer far more insight than would the original intended measurements.

A.3 Convolutions

The sum of a pair of random variables is itself a random variable. Consider the sum $z = x + y$, for random variables x distributed as density $X(x)$, and y distributed as density $Y(y)$. One may express distribution (density) $Z(z)$

for random variable z, and define the convolution by:

$$Z(z) = (X * Y)(z) = \int dt\, X(t) \cdot Y(z-t) = \int dt\, Y(t) \cdot X(z-t) \quad \text{(A.25)}$$

The integration generally poses complications in this spatial view. One may however simplify, via Fourier or Laplace transform, defined respectively as:

$$\mathcal{F}_{x\to k}\{f(x)\}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dx\, e^{-ikx}\, f(x) \quad \text{(A.26)}$$

$$\mathcal{L}_{x\to s}\{f(x)\}(s) = \int_{0}^{\infty} dx\, e^{-sx}\, f(x) \quad \text{(A.27)}$$

If one applies a Fourier transform to a general density over $\mathcal{R}$ or a Laplace transform to a density limited to $\mathcal{R}^+$, the convolution is replaced by simple scalar multiplication:

$$\mathcal{F}_{z\to k}\{(X * Y)(z)\}(k) = \mathcal{F}_{x\to k}\{X(x)\}(k) \cdot \mathcal{F}_{y\to k}\{Y(y)\}(k) \quad \text{(A.28)}$$

$$\mathcal{L}_{x\to s}\{(X * Y)(z)\}(s) = \mathcal{L}_{x\to s}\{X(x)\}(s) \cdot \mathcal{L}_{y\to s}\{Y(y)\}(s) \quad \text{(A.29)}$$

By inference from the transforms, the convolution satisfies associativity, distributivity and commutivity, with respective expressions:

$$[(X * Y) * Z](t) = [X * (Y * Z)](t) \quad \text{(A.30)}$$

$$[(aX + bY) * Z](t) = a(X * Z)(t) + b(Y * Z)(t) \quad \text{(A.31)}$$

$$(X * Y)(t) = (Y * X)(t) \quad \text{(A.32)}$$

A few particular results are of note, with regard to the research at hand. Viewed in transform, one gains insight into the structure of convolution, and into composite probabalistic phenomena studied throughout this work.

The Fourier transform (equation A.26) of the Gaussian density (equation A.12) may be written as follows:

$$\mathcal{F}_{x\to k}\{g(x;\mu,\sigma)\}(k) \;=\; e^{-ik\cdot\mu}\cdot e^{-\frac{1}{2}\sigma^2k^2} \qquad \text{(A.33)}$$

One deduces that convolutions amongst Gaussian densities yield Gaussian results, by following a simple rule:

$$(g(;\mu_a,\sigma_a) * g(;\mu_b,\sigma_b))(x) \;=\; g(x;\mu=\mu_a+\mu_b,\,\sigma=\sqrt{\sigma_a^2+\sigma_b^2}) \qquad \text{(A.34)}$$

For functions which are nearly Gaussian, convolutions approach Gaussian form, in some sense even more closely. Theoretical investigation of the phenomenon yields the *central limit theorem*. Some related notes may be found in Appendix A.4.

Another example yields a more richly structured result. Laplace transform (equation A.27) of the general gamma density (equation A.13) yields a simple expression:

$$\mathcal{L}_{x\to s}\{\gamma(x;\alpha,\beta)\}(s) \;=\; (1+\beta s)^{-\alpha} \qquad \text{(A.35)}$$

With view to equation A.35, one may infer the following set of exact theoretical results pertaining to gamma density convolutions:

$$(\gamma(;\alpha_A,\beta) * \gamma(;\alpha_B,\beta))(x) \;=\; \gamma(x;\alpha_A+\alpha_B,\beta) \qquad \text{(A.36)}$$

$$(\prod_k (*)\gamma(;1;\beta_k))(x) \;=\; \sum_k g_k\gamma(x;1,\beta_k); \qquad \text{(A.37)}$$

$$\text{where}\quad g_k=\prod_{\substack{l\\ l\neq k}}\left(\frac{1}{1-\frac{\beta_l}{\beta_k}}\right) \;\Leftrightarrow\; k\neq l\to\beta_k\neq\beta_l$$

$$\text{if} \quad f(x) = \sum_k F_k \gamma(x; \alpha_{fk}, \beta_{fk}) \quad \text{where} \quad \sum_k F_k = 1, \tag{A.38}$$

$$\text{and} \quad g(x) = \sum_k G_k \gamma(x; \alpha_{gk}, \beta_{gk}) \quad \text{where} \quad \sum_k G_k = 1,$$

$$\text{with} \quad \alpha_{fk}, \alpha_{gk} \epsilon \, \mathbf{N} \quad \text{and} \quad h(x) = (f() * g())(x) \quad \text{then}$$

$$h(x) = \sum_k H_k \gamma(x; \alpha_{hk}, \beta_{hk}) \quad \text{where} \quad \sum_k H_k = 1$$

Formulas A.36 through A.38 show that the span of such gamma functions closes under convolution. The results are theoretically elegant; they are also unfortunately impractical from an experimental perspective. One cannot in practice expect to measure such a multitude of free parameters.

One may however acquire a more practical result. Real signals should correspond to multiple convolutions of base gamma signals with various α and β parameter definitions. Composite signals may be expressed practically as asymptotic forms of the same gamma functions, but now without restriction to integral α parameters:

$$(\prod_k (*)\gamma(; \alpha_k; \beta_k))(x) \sim \gamma(x; \alpha = \frac{(\sum \alpha_k \beta_k)^2}{\sum \alpha_k \beta_k^2}, \beta = \frac{\sum \alpha_k \beta_k^2}{\sum \alpha_k \beta_k}) \tag{A.39}$$

The result obtains from second order expansion of equation A.35. Such expression preserves a measurable link between sample mean and standard deviation. One may test real experimental data with this link, for agreement with the hypothesis of gamma distribution.

A.4 Central limit theorem

The concept of convolution may be carried to a sensible limit. One might well ask whether, subject to some normalization, the limit of a highly convolved variable is well defined. The well established answer to this question is in the affirmative.[MATH 70][PATH 72][SVOZ 93][TAYL 98] Detailed examinations [CHAN 89][ITÔ 74][SEN 93][STRA 63] of the *central limit theorem* may be found in the literature. A simple demonstration of the result may be set out as in the following.

Before introducing the central limit theorem, one may establish some simple results for low order moments of convolved distributions. Expansion of the Fourier transform (equation A.26) for general density function $p(x)$ with mean μ and standard deviation σ yields a simple low order expansion:

$$\tilde{p}(k) = \mathcal{F}_{x\to k}\{p(x)\}(k) \sim 1 + \mu\cdot(ik) + \frac{1}{2}(\mu^2+\sigma^2)\cdot(ik)^2 + \ldots \quad \text{(A.40)}$$

Composite function $x = \sum x_n$ (with x_n independent) has convolved probability density $P(x) = (\prod(*)p_n())(x)$. Fourier transformation with application of equation A.28 yields following low order expansion:

$$\begin{aligned} 1 \;+\; & \mu\cdot(ik) \;+\; \frac{1}{2}(\mu^2+\sigma^2)\cdot(ik)^2 \;+\; \ldots \;\sim \quad \text{(A.41)} \\ 1 \;+\; & \textstyle\sum\mu_n\cdot(ik) \;+\; \frac{1}{2}[(\sum\mu_n)^2+\sum\sigma_n^2]\cdot(ik)^2 \;+\; \ldots \end{aligned}$$

One finds, equating terms order by order, the following general results:

$$\begin{aligned} \mu &= \textstyle\sum\mu_n \quad \text{(A.42)} \\ \sigma^2 &= \textstyle\sum\sigma_n^2 \end{aligned}$$

These results may also be derived without resort to Fourier transformation. Didactically, however, one infers that the relationship between observed and transformed views can illuminate key concepts of convolution.

With view to the convolution scaling inferred from equation A.42, one may set up the conventional form of the central limit theorem. One first simplifies the multiple convolution by setting all included probability densities identically, so that $p_n(x) = p(x)$ with mean μ and variance σ^2. One further defines related density $q_N(z)$ as follows:

$$p(x) = \frac{\sqrt{N}}{\sigma} \cdot q_N(\frac{\sqrt{N}}{\sigma}(x-\mu)) \tag{A.43}$$

By construction, distribution $q_N(z)$ has mean 0 and variance $1/N$. More completely, one may write the following Fourier transforms for such densities:

$$\begin{aligned} \tilde{q}_N(k) &= \mathcal{F}_{x\to k}\{q_N(x)\}(k) = 1 - \frac{1}{2}\cdot(\frac{k}{\sqrt{N}})^2 + (\frac{k}{\sqrt{N}})^3 \cdot \epsilon(\frac{k}{\sqrt{N}}) \qquad \text{(A.44)} \\ \tilde{p}(k) &= \mathcal{F}_{x\to k}\{p(x)\}(k) = e^{-ik\mu}\cdot[1 - \frac{1}{2}\cdot(\sigma k)^2 + (\sigma k)^3 \cdot \epsilon(\sigma k)] \end{aligned}$$

Here, one assumes that function $\epsilon(z)$ be analytic about $z = 0$. One finds concurrently, by transforming the integration variable in equation A.25, the multiple convolution identity:

$$P_N(x) = (\prod_{n=1}^{n=N}(*)_x p())(x) = \frac{1}{\sqrt{N}\,\sigma}(\prod_{n=1}^{n=N}(*)_z q_N())(z = \frac{x - N\mu}{\sqrt{N}\,\sigma}) \tag{A.45}$$

Proof of the central limit theorem now derives naturally as follows. One extracts normalized density $Q_N(z)$ with mean 0 and variance 1, as a multiple

convolution with respect to its parameter space:

$$Q_N(z) \;=\; (\prod_{n=1}^{n=N}(*)q_N())(z) \tag{A.46}$$

Viewed in Fourier transform, $Q_N(z)$ appears as a simple asymptotic form:

$$\begin{aligned}
\tilde{Q}_N(k) \;&=\; \prod_{n=1}^{n=N}\tilde{q}_N(k) \;=\; \tilde{q}_N^N(k) \qquad (A.47)\\
&=\; \{1-\frac{1}{2}\cdot(\frac{k}{\sqrt{N}})^2+(\frac{k}{\sqrt{N}})^3\cdot\epsilon(\frac{k}{\sqrt{N}})\}^N\\
&\sim\; \exp(-\frac{1}{2}k^2+\frac{k^3}{\sqrt{N}}\cdot\epsilon(\frac{k}{\sqrt{N}}))\\
&\sim\; \tilde{Q}(k) \;=\; \exp(-\frac{1}{2}k^2)
\end{aligned}$$

The asymptotic distribution $Q(z)$ appears, via equation A.33, as the central limit (standard normal) Gaussian with mean $\mu = 0$ and variance $\sigma^2 = 1$:

$$Q(z) \;=\; g(\,z;\,\mu=0,\,\sigma=1\,) \;=\; \frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}z^2} \tag{A.48}$$

Rigorous derivation would require investigation of conditions *(Lindeberg-Feller)* under which a density actually converges to the asymptotic central limit. Many excellent texts study that issue.[FELL 68][ITÔ 74][STRA 63] For purposes of this work, it suffices to note that the *central limit theorem* fails only with densities which are in some sense pathologically constructed.

Example densities may be classified by normalized convolution limit. The densities of equations A.11, A.12 and A.13 convolve toward the *central limit*. The density of equation 2.22, on the other hand fails to converge: that density is indeed pathological, having unbounded variance $\langle(x-\mu)^2\rangle$, unbounded

absolute deviation $\langle |x - \mu| \rangle$, and even non-existent mean. For such *Cauchy* distributions, the central value, median or mode, μ derives only as a principal value, and no moment of higher order converges.

By design, real experimental distributions should observe sensible behaviour subject to strict bounds. One generally expects experimental errors to follow the *central limit*; measurements are in fact often modelled with Gaussian error terms.

The limiting mathematics invoked in derivation of the *central limit theorem* may be generalized into a notational tool, for study of phenomena, where normalized sums of random variables approach the limits of convolution. Such a view is presented in the next appendix.

A.5 Simple stochastic process sums

A modern approach to probability with highly convolved functions involves generalization of concepts developed in the previous appendices A.3 and A.4.[CHUN 83][EMER 89][ITÔ 74][STRA 63] For the sake of simplicity, a rigorous theory of stochastic variables is avoided in the present treatment, with only those tools material to the current research being demonstrated. Stochastic process concepts required for this thesis may be restricted to the simplest form of the general theory.[CHAN 89][FELD 96][SEN 93][TAYL 98] Throughout this work, *simple stochastic process* refers only to those processes modelled by independent increments in continuous-space and contin-

uous time. All *stochastic process sums* are viewed in normalized form.

The fundamental concept of a stochastic process may be viewed as a limiting phenomenon. The classic case is that of the "drunkards walk" envisioned by Albert Einstein as a theoretical model for *Brownian motion.*[CHAN 89] In that example, the diffusion of a particle follows a convolved path of independent steps punctuated by random interactions, and yielding a final distribution. One presumes the limit where interaction number exceeds all bounds. Solution derives from the *central limit theorem* of the previous section.

One may infer a pair of useful generalizations from the case of *Brownian motion.* First, one notes that the process results in a simple probability density or distribution. Physical labels of intermediates, for example timing or location of interactions, do not appear in the final expression. Second, one notes that the concept of process ordering, expressed by a single continuous (time-like) parameter is not essential in realization of such a path. One may rearrange the process order at will without altering the total realized probability density or distribution. The *simple stochastic process sum* may thus be viewed as a conceptual model including more general stochastic processes.

Detailed treatment of the subject would at this point require precise definitions for *Brownian motion* and *Wiener process*, and as well more general stochastic methods with the derivation of *Itô's equation.*[ITÔ 74] Present requirements limit the scope to demonstration of notational tools for basic stochastic calculations. Precise development of the tools employed here may be found in many excellent treatises on the subject of stochastic variables.

The simple stochastic process sum currently viewed may be modelled as follows. A set of limiting independent continuous-space increments is summed over a continuous-time parameter. The time-like parameter should be understood to address all contributions in some unprescribed order. Since reordering leaves net results unaffected, precise contruction of time-like intervals may be avoided. The resulting simple stochastic process may be summed notationally by writing the net stochastic variable A in terms of microscopic limiting steps $dA(s)$ as follows:

$$A = \int_s dA(s) \tag{A.49}$$

One may extend the concepts of expectation expressed in Appendix A.1 to stochastic variables in the same notation:

$$\langle A \rangle = \int_s \langle dA(s) \rangle \tag{A.50}$$

One may express more general expectations by functional expansion. The concept of independent local contributions arises naturally as follows:

$$\langle dA(s')\, dA(s'') \rangle = \langle dA(s') \rangle \langle dA(s'') \rangle + \delta(s', s'') \cdot d\sigma^2_{dA(s')\, dA(s'')} \tag{A.51}$$

The term with the (pseudo) Dirac delta-function $\delta(s', s'')$ identifies the point auto-correlation. Continuing with the notational tool one may write:

$$\langle A \cdot A \rangle = \langle \int_{s'} \int_{s''} dA(s')\, dA(s'') \rangle = \langle A \rangle_{s'} \langle A \rangle_{s''} + \langle \int_s d\sigma^2_{dA(s)\, dA(s)} \rangle_s \tag{A.52}$$

One has thus recovered for stochastic processes the variance expansion of equation A.6, by drawing the following association:

$$\sigma^2_{AA} = \langle \int_s d\sigma^2_{dA(s)\, dA(s)} \rangle_s \longrightarrow \langle A \cdot A \rangle = \langle A \rangle^2 + \sigma^2_{AA} \tag{A.53}$$

Higher moments may be identified by combinatorics on algebraic expressions. For example, letting $\bar{A} = \langle A \rangle$, one may examine integral powers for the stochastic variable $A - \bar{A}$. While expectations for (antisymmetric) odd powers vanish, (symmetric) even powers yield the following general result:

$$\langle (A - \bar{A})^{2N} \rangle = \frac{(2N)!}{2^N} \sigma_{AA}^{2\,N} \tag{A.54}$$

One quickly recognizes the combinatorics for this simple stochastic variable as identically those of the well known *Wick expansion* for real scalar quantum fields. The approach of *stochastic quantization* explores particle physics from exactly this point of view. The interested reader is referred to the literature.[GLIM 81] One may moreover infer a concrete probability density associated with stochastic variable A, by noting that the moments of equation A.54 match identically with the Gaussian form:

$$P(A) = \frac{1}{\sqrt{2\pi\sigma_{AA}^2}} e^{-\frac{(A-\bar{A})^2}{2\sigma_{AA}^2}} \tag{A.55}$$

Wiener integration confirms recovery of the *central limit* as net result for such a (limiting) stochastic process.

For the thesis at hand, one requires a minor extension of the above scalar Gaussian example. Consider here a complex stochastic variable where the real and complex components A_x & A_y satisfy respectively:

$$\langle dA_x(s')\, dA_x(s'') \rangle = \langle dA_x(s') \rangle \langle dA_x(s'') \rangle + \delta(s', s'') \cdot \tfrac{1}{2} d\sigma^2_{dA(s')\, dA(s'')} \tag{A.56}$$

$$\langle dA_y(s')\, dA_y(s'') \rangle = \langle dA_y(s') \rangle \langle dA_y(s'') \rangle + \delta(s', s'') \cdot \tfrac{1}{2} d\sigma^2_{dA(s')\, dA(s'')}$$

$$\langle dA_x(s')\, dA_y(s'') \rangle = \langle dA_x(s') \rangle \langle dA_y(s'') \rangle$$

One finds that expression via complex stochastic variable A and conjugate A^* simplifies combinatoric algebra for salient functional expressions:

$$A = A_x + iA_y \quad \& \quad A^* = A_x - iA_y \quad \longrightarrow \tag{A.57}$$

$$\begin{aligned}
\langle d(A-\bar{A})(s')\, d(A^*-\bar{A}^*)(s'')\rangle &= \delta(s',s'')\cdot d\sigma^2_{dA(s')\,dA(s'')} \\
\langle d(A-\bar{A})(s')\, d(A-\bar{A})(s'')\rangle &= 0 \\
\langle d(A^*-\bar{A}^*)(s')\, d(A^*-\bar{A}^*)(s'')\rangle &= 0
\end{aligned}$$

Consider for example stochastic variable I, based on A & A^*, whose moments and expectations may be expressed combinatorically as follows:

$$I = (A-\bar{A})(A^*-\bar{A}^*) \quad \longrightarrow \quad \langle I^N\rangle = N!\cdot {\sigma^2_{AA}}^N \tag{A.58}$$

One may expand the Laplace transform for the related density function:

$$\mathcal{L}_{I\to s}\{P(I)\}(s) = \sum_{n=0}^{\infty}\frac{1}{n!}\langle I^n\rangle s^n = \sum_{n=0}^{\infty}(\sigma^2_{AA}s)^n = \frac{1}{1+\sigma^2_{AA}s} \tag{A.59}$$

$$P(I) = \gamma(I;\,\alpha=1,\,\beta=\sigma^2_{AA})$$

The expression corresponds precisely to the Laplace transform of the exponential density $\gamma(I;1,\,\sigma^2_{AA})$, where the general gamma density follows equation A.13. One may confirm this by comparing with general gamma density Laplace transform of equation A.35. Higher order convolutions may be evaluated by stochastic methods, or equivalently by the traditional methods of Appendix A.3.

A.6 Stochastic EM waves

One may apply stochastic methods presented in Appendix A.5 as notational tools for derivation of some physically meaningful expressions. While traditional statistical methods do yield identical results, the older methods also require repetitive reduction of stochastic limits. The newer stochastic approach best demonstrates its utility by example.

A.6.1 Scalar wave approximation

Consider the simple forward propagating complex plane wave in a single spatial dimension:

$$A(z,t) = A\,e^{+i(kz-\omega t)} \quad \& \quad I_{\pm}(z,t) = \pm\frac{1}{2}\kappa\,(A(z,t) \pm A^{*}(z,t))^{2} \qquad \text{(A.60)}$$

One notes that complex scalar $A = |A|e^{+i\phi}$ and conjugate $A^{*} = |A|e^{-i\phi}$ parametrize both magnitude and phase offset. Instantaneous instensities $I_{\pm}(z,t)$ derive conventionally from the respective flow rates of energy content in the real ($\Re\{w\} = \frac{1}{2}(w + w^{*})$) or imaginary ($\Im\{w\} = \frac{1}{2i}(w - w^{*})$) wave components, with choice of constant κ deriving from intensity normalization.

One often measures wave-related quantities in time-average rather than instantaneously. For the simple wave of equation A.60, one notes the following time-averages:

$$\langle A(z,t)\rangle_{t} = 0 \quad \longrightarrow \quad \langle I_{\pm}(z,t)\rangle_{t} = \kappa AA^{*} \qquad \text{(A.61)}$$

One notes that time-averaging leaves unaffected the constants A & A^{*}, and

thereby $|A|$ & ϕ, defining the original instantaneous wave $A(z,t)$. A related inference follows, if one views constants A & A^* as random variables, and views time sampling as similarly randomizing. Expectations over time and expectations over wave parameters may be expressed as a single process:

$$\langle\, \langle\, \Phi(z,t)\rangle_t \rangle_{AA^*} \;=\; \langle \Phi(z,t)\rangle \;=\; \langle\, \langle\, \Phi(z,t)\rangle_{AA^*} \rangle_t \qquad \text{(A.62)}$$

With this insight, one realizes the potential of stochastic methods for waves of such description.

As preliminary example of stochastic plane waves, consider the theoretical (time-averaged) intensity distribution for $i = \langle I_{\pm}(z,t)\rangle_t$. Examination of the distribution for related random variable AA^* appears in the preceding Appendix A.5. One infers from expressions of equations A.56 through A.59 the distribution of (time-averaged) intensity:

$$P(i) = \frac{1}{\langle i \rangle} \exp(-\frac{i}{\langle i \rangle}) \quad \text{where} \quad i = \langle I_{\pm}(z,t)\rangle_t \qquad \text{(A.63)}$$

Thus recovers the intensity distribution of *Goodman*,[GOOD 89] whose result also appears quoted in equation 1.7.

One may derive other results for plane waves, (as per equation A.60,) in single spatial dimension. Consider introduction of phase angle offset θ, denoted by subscript:

$$A_\theta = A\,e^{+i\theta} \;\longrightarrow\; A_\theta(z,t) = e^{+i\theta}\,A\,(z,t) = A\,e^{+i(kz-\omega t+\theta)} \quad \text{(A.64)}$$
$$A^*_\theta = A^* e^{-i\theta} \;\longrightarrow\; A^*_\theta(z,t) = e^{-i\theta} A^*(z,t) = A^* e^{-i(kz-\omega t+\theta)}$$

Time-average intensity shows trivial invariance $\langle I_\theta \rangle_t = \langle I_0 \rangle_t$ under phase offset. One therefore cannot perform direct measurement of phase offset for a single time-average intensity.

One may however measure such an offset indirectly against a reference signal. Consider a simple pair of complex waves, $A'(z,t)$ & $A''(z,t)$, of the type introduced in equation A.60. The pair is coherent so long as $\omega' = \omega''$. Interference of the wave pair obtains by simple addition of complex amplitudes. Symbolically, one may effect interference in (complex) wave amplitude and (real) time-average intensity, respectively as follows:

$$A = A' + A'' \quad \& \quad A^* = A'^* + A''^* \quad \longrightarrow \tag{A.65}$$

$$\langle I \rangle_t = \langle I' \rangle_t + \langle I'' \rangle_t + \sqrt{\langle I' \rangle_t \langle I'' \rangle_t} \cos(\phi' - \phi'')$$

The result allows for simple estimation of phase differences, provided one may rely on direct accurate measurements of the contributing intensities. Experimental measurement of these poses serious logistical complications. One naturally asks whether a related estimation may be constructed, which relies still less directly on the contributing intensities.

One may construct a requisite measure of phase difference in the form of a nontrivial statistical observable. Regarding wave components as deriving from independent coherent stochastic sources, one infers an independent pair of complex stochastic variables, each following equation A.57, with vanishing expectations $\langle A' \rangle = 0$ & $\langle A'' \rangle = 0$. Independence of the interfering stochastic

variable pair imposes the following explicit expectations:

$$\begin{aligned}
\langle d(A'-\bar{A}')(s')\, d(A''-\bar{A}'')(s'')\rangle &= 0 \qquad \text{(A.66)}\\
\langle d(A'-\bar{A}')(s')\, d(A''^{*}-\bar{A}''^{*})(s'')\rangle &= 0\\
\langle d(A'^{*}-\bar{A}'^{*})(s')\, d(A''-\bar{A}'')(s'')\rangle &= 0\\
\langle d(A'^{*}-\bar{A}'^{*})(s')\, d(A''^{*}-\bar{A}''^{*})(s'')\rangle &= 0
\end{aligned}$$

One finds, surprisingly, that measurement of only time-average net intensities suffices. Consider low order moments of time-average intensities under phase shift as outlined by equation A.64. If A' and A'' are phase offset by angles α and β respectively, intensity moments calculate simply:

$$\langle I_{\alpha\beta}\rangle = \langle I_{00}\rangle = \langle I\rangle = \langle I'\rangle + \langle I''\rangle \quad \text{and} \qquad \text{(A.67)}$$

$$\begin{aligned}
\langle\, \langle I_{\alpha\beta}\rangle_t \langle I_{00}\rangle_t \,\rangle &= \kappa^2 \,\langle\, (A'_\alpha + A''_\beta)(A'^{*}_\alpha + A''^{*}_\beta)\,(A'_0 + A''_0)(A'^{*}_0 + A''^{*}_0)\,\rangle\\
&= \kappa^2 \,\langle\, (A'_\alpha A'^{*}_\alpha + A''_\beta A''^{*}_\beta)(A'_0 A'^{*}_0 + A''_0 A''^{*}_0)\,\rangle\\
&+ \kappa^2 \,\langle\, A'_\alpha A''^{*}_\beta A''_0 A'^{*}_0 + A''_\beta A'^{*}_\alpha A'_0 A''^{*}_0 \,\rangle\\
&= \kappa^2 \,\langle\, A'A'^{*}A'A'^{*} + A''A''^{*}A''A''^{*} \,\rangle\\
&+ 2\,[1+\cos(\alpha-\beta)]\,\kappa^2 \,\langle\, A'A'^{*}A''A''^{*} \,\rangle\\
&= 2\,\{\,\langle I'\rangle^2 + [1+\cos(\alpha-\beta)]\langle I'\rangle\langle I''\rangle + \langle I''\rangle^2\,\}
\end{aligned}$$

Intensity covariance under phase shift derives naturally from the results of equation A.67, which may be expressed as follows:

$$\langle\, \langle I_{\alpha\beta}-\bar{I}_{\alpha\beta}\rangle_t \langle I_{00}-\bar{I}_{00}\rangle_t \,\rangle = \langle I'\rangle^2 + 2\cos(\alpha-\beta)\langle I'\rangle\langle I''\rangle + \langle I''\rangle^2 \qquad \text{(A.68)}$$

One may write the corresponding standard Pearson's cross-correlation:

$$\rho(\langle I_{\alpha\beta}\rangle_t, \langle I_{00}\rangle_t) = \frac{\langle \langle I_{\alpha\beta} - \bar{I}_{\alpha\beta}\rangle_t \langle I_{00} - \bar{I}_{00}\rangle_t \rangle}{\sqrt{\langle \langle I_{\alpha\beta} - \bar{I}_{\alpha\beta}\rangle_t^2 \rangle \langle \langle I_{00} - \bar{I}_{00}\rangle_t^2 \rangle}} \tag{A.69}$$

Evaluation of the cross-correlation of equation A.69, respecting phase shifted plane wave covariances expressed by equation A.68, yields a final form:

$$\begin{aligned} \rho(\langle I_{\alpha\beta}\rangle_t, \langle I_{00}\rangle_t) &= 1 + \kappa(\langle I'\rangle, \langle I''\rangle)\,[\cos(\alpha - \beta) - 1] \\ \kappa(\langle I'\rangle, \langle I''\rangle) &= \frac{2\langle I'\rangle \langle I''\rangle}{(\langle I'\rangle + \langle I''\rangle)^2} \end{aligned} \tag{A.70}$$

One has thus recovered the simple result of *Goodman*, which appears quoted in section 1.2. One may note that such identity holds only subject to some rather severe limitations. The wave of the foregoing derivation, as defined in equation A.60, has been simplified to a complex plane wave of single spatial dimension and direction. In the case of general electromagnetism, the result holds only as scalar approximation to the true phenomenon.

One may however draw an inference for extrapolation of equation A.70 to realistic waves. Trivial phase shift invariance, $\langle I^N_{\alpha\beta}\rangle = \langle I^N_{00}\rangle$, holds for all moments of a single measured intensity. Pearson's cross-correlation (per equation A.69) thus simplifies:

$$\rho(\langle I_{\alpha\beta}\rangle_t, \langle I_{00}\rangle_t) = \frac{\langle \langle I_{\alpha\beta} - \bar{I}_{\alpha\beta}\rangle_t \langle I_{00} - \bar{I}_{00}\rangle_t \rangle}{\langle \langle I_{00} - \bar{I}_{00}\rangle_t \langle I_{00} - \bar{I}_{00}\rangle_t \rangle} \tag{A.71}$$

One observes for cross-correlations respecting equation A.71, that the denominator derives generally as limiting case (ie. $(\alpha - \beta) \rightarrow 0$) of the numerator. Situations, where the cross-moment takes a particularly simple

form, yield cross-correlations of well defined description:

$$\langle \langle I_{\alpha\beta} - \bar{I}_{\alpha\beta} \rangle_t \langle I_{00} - \bar{I}_{00} \rangle_t \rangle = S \cdot \cos(\alpha - \beta) + T; \quad S\&T \text{ free of } \alpha\&\beta \quad \text{(A.72)}$$

$$\longrightarrow \quad \rho(\langle I_{\alpha\beta} \rangle_t, \langle I_{00} \rangle_t) = 1 + \kappa \cdot [\cos(\alpha - \beta) - 1] \quad \text{where } \kappa = (1 + \frac{T}{S})^{-1}$$

Thus recovers, for any phase-shifted intensity cross-correlation satisfying equation A.72, precisely the theoretical format requisite to investigations of this thesis. While real solutions for (function) $\kappa(\ldots)$ may vary, functional integrity of the cross-correlation remains transfixed.

A.6.2 Electromagnetic plane waves

One may apply stochastic methods to experiments involving general electromagnetic plane waves. In line with experimental concepts of laser speckle interferometry, as presented in this thesis, one restricts study to stochastic light sources of a very particular geometry. One should like, nonetheless, to deal only with experimentally realistic solutions to the Maxwell equations.

The basic source term for the multidimensional case may be written as asymptotic to the plane wave vector potential. In Coulomb gauge, such a vector potential, and its derivatives may be expressed as follows:[JACK 62]

$$\vec{A}(\vec{r},t) = \vec{A}\, e^{i(\vec{k}\cdot\vec{r}-\omega t)} \quad \& \quad 0 = \vec{\nabla} \cdot \vec{A}(\vec{r},t) = \vec{k} \cdot \vec{A}(\vec{r},t) \qquad \text{(A.73)}$$

$$\vec{E}(\vec{r},t) = -\frac{1}{c}\partial_t \vec{A}(\vec{r},t) \quad \& \quad \vec{B}(\vec{r},t) = \vec{\nabla} \times \vec{A}(\vec{r},t) \quad \longrightarrow$$

$$\vec{E}(\vec{r},t) = i\frac{\omega}{c}\vec{A}(\vec{r},t) \quad \& \quad \vec{B}(\vec{r},t) = i\vec{k} \times \vec{A}(\vec{r},t)$$

One requires electromagnetic ($\vec{E}$ & $\vec{B}$) fields to comply with free space Maxwell equations. To this end, a pair of obvious constraints may be imposed. First, real fields must observe $\tilde{\nabla} \cdot \vec{A}(\vec{r},t) = 0$, the Coulomb gauge constraint. Second, simple waves follow $c = f \cdot \lambda$ as standard wave velocity identity, with light speed c, frequency $f = 2\pi\omega$ and wavelength $\lambda = 2\pi/|\tilde{k}|$. For plane waves of the type considered, such constraints read respectively:

$$\tilde{k} \cdot \vec{A} = 0 \quad \& \quad \omega = |\tilde{k}|\, c \tag{A.74}$$

Subject to equations A.73 and A.74, one may quickly confirm the free space Maxwell equations for simple plane waves of such description:

$$\begin{aligned} \tilde{\nabla} \cdot \vec{E}(\vec{r},t) &= 0 \qquad \vec{\nabla} \times \vec{E}(\vec{r},t) + \frac{1}{c}\partial_t \vec{B}(\vec{r},t) = \vec{0} \\ \tilde{\nabla} \cdot \vec{B}(\vec{r},t) &= 0 \qquad \vec{\nabla} \times \vec{B}(\vec{r},t) - \frac{1}{c}\partial_t \vec{E}(\vec{r},t) = \vec{0} \end{aligned} \tag{A.75}$$

Light intensity in reality derives from energy propagation impinging upon a surface at some incident angle, to be observed with some proportionality constant σ. One may write this result, up to a pure boundary term, in terms of the Poynting vector $\vec{S}(\vec{r},t)$:[JACK 62]

$$\vec{S}(\vec{r},t) = 2\sigma\, \vec{E}(\vec{r},t) \times \vec{B}(\vec{r},t) \quad \longrightarrow \quad I(\vec{r},t;\hat{n}) = \vec{S}(\vec{r},t) \cdot \hat{n} \tag{A.76}$$

In view of the current complex vector field notation, set out in equation A.73, one further defines Poynting vector for real ($\Re\{w\} = \frac{1}{2}(w + w^*)$) or imaginary ($\Im\{w\} = \frac{1}{2i}(w - w^*)$) wave components respectively as:

$$\vec{S}_{\pm}(\vec{r},t) = \pm\frac{1}{2}\sigma\,(\vec{E} \pm \vec{E}^*)(\vec{r},t) \times (\vec{B} \pm \vec{B}^*)(\vec{r},t) \tag{A.77}$$

Time-average views of the Poynting vector and intensity derive from scalar multiplication of independent terms. Potential constant $\vec{A}$ and conjugate $\vec{A}^*$ give the time-average flow rate of energy content. Propagation ray, defined by unit $\hat{k} = \frac{\vec{k}}{|\vec{k}|}$, and intersecting surface with unit normal $\hat{n}$, define the incidence. One may write, for waves respecting equation A.73, following simple time-average observables:

$$\langle \vec{S}_{\pm} \rangle_t = \sigma (\frac{\omega}{c})^2 (\vec{A} \cdot \vec{A}^*) \hat{k} \quad \longrightarrow \quad \langle I_{\pm}(;\hat{n}) \rangle_t = \langle \vec{S}_{\pm} \rangle_t \cdot \hat{n} \qquad \text{(A.78)}$$

One must note that while wave solutions do propagate in free space, reality nonetheless imposes conditions at some boundary. There, some set of oscillating multipole or surface current terms may be imposed, to model real boundary conditions. In any case, driving oscillatory source $\vec{D}(\vec{r},t)$ amounts to complex constant (tensor) muliplication on the radiating vector potential: symbollically, $\vec{D}(\vec{r},t) \propto \vec{A}(\vec{r},t)$ at the boundary.

One may draw an inference relating stochastic fields and boundary conditions. Stochasticity of either boundary sources $\vec{D}$ or radiating fields $\vec{A}$ imposes stochasticity on both. The result may be applied to stochastic (asymptotic) plane waves following equation A.73 in form. General (coherent) stochastic sources derive from viewing complex vector potential constant $\vec{A}$ and conjugate $\vec{A}^*$ as stochastic sums. Such a view affords practical utility: stochastic effects may be elegantly modelled in complex vector algebra.

One does rely upon verisimilitude of the asymptotic plane wave inference, in the study of speckle interferometry. Validity of the inference may be

shown by considering a single micro-source of a localized composite source. The radiation field may be technically expanded by solution of the Helmholz equation. Proper treatment requires expansion in vector harmonics, with all transverse $\vec{E}$ and transverse $\vec{B}$ modes represented.[JACK 62] Interferometry is however restricted to the far field, where solutions simplify asymptotically:

$$\vec{A}(\vec{r},t) \sim \frac{1}{|\vec{r}|}\vec{A}(\Omega=(\theta,\phi))\, e^{i(\vec{k}\cdot\vec{r}-\omega t)} \tag{A.79}$$

Restricting observation to a region where variation of inverse radius $\frac{1}{|\vec{r}|}$ and angular variation of $\vec{A}(\Omega)$ seem negligible, one extracts asymptotically the requisite coherent plane wave. Viewed from precisely such an observation region, micro-sources of a localized source may be combined as coherent, provided the angular distribution of the net source remains negligible, when viewed from the respective observation point.

A.6.3 Laser speckle intensity distribution

A single asymptotic stochastic source may be utilized to model the expected intensity distribution due to a single real scattering source. One may in fact almost immediately write out the result, which appears quoted in subsection 2.1.1. The obvious stochastic (micro-source) tensor expectations may be expressed as follows:

$$\begin{aligned} \langle d\vec{A}(s)\rangle &= \vec{0} \qquad \langle d\vec{A}(s')\otimes d\vec{A}(s'')\rangle = \vec{\vec{0}} \\ \langle d\vec{A}^*(s)\rangle &= \vec{0} \qquad \langle d\vec{A}^*(s')\otimes d\vec{A}^*(s'')\rangle = \vec{\vec{0}} \end{aligned} \tag{A.80}$$

$$\sigma\,(\frac{\omega}{c})^2\,\langle d\vec{A}(s')\otimes d\vec{A}^*(s'')\rangle = \delta(s',s'')\cdot d\mathcal{I}_{d\vec{A}(s')\,d\vec{A}^*(s'')}$$

One infers construction of observable hermitian intensity tensor $\mathcal{I}$:

$$\mathcal{I} = \int_s d\mathcal{I}_{d\vec{A}(s)\, d\vec{A}^*(s)} \tag{A.81}$$

Micro-sources share an identical (asymptotic) propagation ray. One may therefore infer also a tensor constraint: $\mathcal{I}$ necessarily has singular kernel containing propagation (ray) axis $\hat{k}$:

$$\hat{k} \cdot d\vec{A}(s) = 0 \quad \longrightarrow \quad \hat{k} \cdot \mathcal{I} \cdot \hat{k} = 0 \tag{A.82}$$

With a null eigen-mode corresponding to the (ray) axis of plane wave propagation, orthogonal transformation from principal axes reads as follows:

$$\mathcal{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} = R \begin{pmatrix} I_\perp & 0 & 0 \\ 0 & I_\tau & 0 \\ 0 & 0 & 0 \end{pmatrix} \tilde{R} \tag{A.83}$$

One observes that orthogonal transformation R should sensibly reduce to a real rotation. The complex form corresponds to introduction of elliptical polarization. Random scattering from a rough surface should lack any such phase coherent effects. One deduces that truly random scatter should render tensor $\mathcal{I}$ real and symmetric.

Observable moments of time-average intensity derive simply by application of equations A.76 through A.78. Direct evaluation of stochastic expectations yields the following results:

$$\begin{aligned} (\frac{1}{\hat{k} \cdot \hat{n}}) \langle \langle I_\pm(;\hat{n}) \rangle_t \rangle &= [\sigma (\tfrac{\omega}{c})^2] \langle \vec{A} \cdot \vec{A}^* \rangle = trace(\mathcal{I}) \\ (\frac{1}{\hat{k} \cdot \hat{n}})^2 \langle \langle I_\pm(;\hat{n}) \rangle_t^2 \rangle &= [\sigma (\tfrac{\omega}{c})^2]^2 \langle (\vec{A} \cdot \vec{A}^*)^2 \rangle = trace^2(\mathcal{I}) + trace(\mathcal{I}\mathcal{I}) \end{aligned} \tag{A.84}$$

As particular results, time-average intensity mean and variance values obtain for real scattering from (asymptotic) pointlike stochastic sources:

$$i = \langle I_{\pm}(;\hat{n})\rangle_t = \langle \vec{S}_{\pm}\rangle_t \cdot \hat{n} = I(\hat{k}\cdot\hat{n}) \longrightarrow \tag{A.85}$$

$$\langle i\rangle = trace(\mathcal{I})(\hat{k}\cdot\hat{n}) \quad \& \quad \sigma_{ii}^2 = trace(\mathcal{I}\mathcal{I})(\hat{k}\cdot\hat{n})^2$$

One may finally derive directly the perceived intensity distribution, by noting that choice of axes along principal moments simplifies differential expectations. Most first and second order moments vanish, with only the following pair of exceptions:

$$\kappa\langle dA_{\perp}(s')\, dA_{\perp}^{*}(s'')\rangle = \delta(s'-s'')\cdot di_{\perp}(s',s'') \tag{A.86}$$

$$\kappa\langle dA_{\top}(s')\, dA_{\top}^{*}(s'')\rangle = \delta(s'-s'')\cdot di_{\top}(s',s'')$$

$$\text{where} \quad \kappa = \sigma(\frac{\omega}{c})^2(\hat{k}\cdot\hat{n})$$

One infers that independent components $i_{\perp}$ & $i_{\top}$ are each distributed exponentially, as were simple waves of single spatial dimension, per equation A.63. As a result, observed intensity should be distributed by gamma convolution, per equations A.36 and A.37, as follows:

$$P(i) = \begin{cases} \frac{1}{1-\frac{\langle i_{\top}\rangle}{\langle i_{\perp}\rangle}}\gamma(i;1,\langle i_{\perp}\rangle) + \frac{1}{1-\frac{\langle i_{\perp}\rangle}{\langle i_{\top}\rangle}}\gamma(i;1,\langle i_{\top}\rangle) & \langle i_{\perp}\rangle \neq \langle i_{\top}\rangle \\ \gamma(i;2,\frac{1}{2}(\langle i_{\perp}\rangle + \langle i_{\top}\rangle)) & \langle i_{\perp}\rangle = \langle i_{\top}\rangle \end{cases} \tag{A.87}$$

$$\langle i\rangle = \langle i_{\perp}\rangle + \langle i_{\top}\rangle \quad \text{and}$$

$$\sigma_{ii}^2 = \langle i_{\perp}\rangle^2 + \langle i_{\top}\rangle^2; \quad \text{or}$$

$$\langle i_{\perp}\rangle = \frac{1}{2}(\langle i\rangle + \sqrt{2\sigma_{ii}^2 - \langle i\rangle^2}) \quad \text{and}$$

$$\langle i_{\top}\rangle = \frac{1}{2}(\langle i\rangle - \sqrt{2\sigma_{ii}^2 - \langle i\rangle^2})$$

One notes with interest, that real intensity distributions may be parametrized solely by observation of net intensity, without resort to polarizing filters.

A.6.4 Laser speckle autocorrelation

One may investigate the spatial scale at which asymptotic plane waves from localized coherent sources demonstrate stochasticity. Phenomenology takes an obvious path in derivation of intensity auto-correlation with respect to shifting observation location. In so doing, one finds that perceived speckle dimension derives not from surface nature at the stochastic scattering source, but rather from the electromagnetic field distribution found there. Results of the current examination appear quoted in subsection 2.1.2.

The auto-correlation under study may be rather simply expressed. One may moreover infer properties which such an expectation must sensibly satisfy. Near and far signal relationships of time-average intensity under auto-correlation derive as trivial limits:

$$\langle\, i(\vec{x})\, i(\vec{y})\,\rangle \;=\; \langle\, i(\vec{x})\,\rangle\langle\, i(\vec{y})\,\rangle \;+\; \rho(\, i(\vec{x}),\, i(\vec{y})\,)\cdot\sqrt{\sigma_{ii}^2(\vec{x})\,\sigma_{ii}^2(\vec{y})} \qquad \text{(A.88)}$$

$$\rho(\, i(\vec{z}+\vec{\epsilon}),\, i(\vec{z}-\vec{\epsilon})\,) \xrightarrow{\vec{\epsilon}\to\infty} 0;\quad \langle\, i(\vec{z}+\vec{\epsilon})\, i(\vec{z}-\vec{\epsilon})\,\rangle \xrightarrow{\vec{\epsilon}\to\infty} \langle\, i(\vec{z}+\vec{\epsilon})\,\rangle\langle\, i(\vec{z}-\vec{\epsilon})\,\rangle$$

$$\rho(\, i(\vec{z}+\vec{\epsilon}),\, i(\vec{z}-\vec{\epsilon})\,) \xrightarrow{\vec{\epsilon}\to 0} 1;\quad \langle\, i(\vec{z}+\vec{\epsilon})\, i(\vec{z}-\vec{\epsilon})\,\rangle \xrightarrow{\vec{\epsilon}\to 0} \langle\, i(\vec{z})\, i(\vec{z})\,\rangle$$

As preliminary to derivation of the asymptotic forms for moments appearing in the auto-correlation, (equation A.88,) one briefly revisits the real scattering field viewed (per equation A.79) in the far field. So long as stochastic effects dominate the autocorrelation one may ignore angular variation in

function $f(\Omega)$ and conjugate $f^*(\Omega)$ as negligible. In anticipation of precisely such a result, angular variation inherent to a single scattering micro-source may be ignored.

Required next is a stochastic expression modelling the real scattering source. Measure for such a stochastic process may be conveniently parametrized in real spatial coordinates. One postulates a homogeneous scattering source distributed in real space by imposing trivial source structure. Structure derives naturally from perceived (micro-source) tensor expectations:

$$\langle d\vec{A}(\vec{a})\rangle = \vec{0} \qquad \langle\, d\vec{A}(\vec{a}') \otimes d\vec{A}(\vec{a}'')\,\rangle = \vec{\vec{0}} \qquad \text{(A.89)}$$
$$\langle d\vec{A}^*(\vec{a})\rangle = \vec{0} \qquad \langle d\vec{A}^*(\vec{a}') \otimes d\vec{A}^*(\vec{a}'')\rangle = \vec{\vec{0}}$$

$$\sigma\,(\frac{\omega}{c})^2\,\langle d\vec{A}(\vec{a}') \otimes d\vec{A}^*(\vec{a}'')\rangle \;=\; \delta(\vec{a}'-\vec{a}'')\cdot p(\,\vec{a}=\frac{1}{2}(\vec{a}'+\vec{a}'')\,)\cdot \mathcal{I}\; da_x da_y da_z$$
$$\text{with constant tensor } \mathcal{I} \quad \text{and normalization } \int da_x da_y da_z\, p(\vec{a}) \;=\; 1$$

Evaluation requires an asymptotic expression for optical path variation under shifting location. One parametrizes location inside the source by vector $\vec{a}$ about scattering centroid $\vec{a}_0$ and defines deviation $\delta\vec{a} = \vec{a} - \vec{a}_0$; observation in the viewing surface, near location $\vec{z}_0$ similarly follows $\delta\vec{z} = \vec{z} - \vec{z}_0$. The intended limiting condition, $\lambda \ll |\delta a| \ll r_{a_0 z_0}$, indicates how variation of equation A.79 might be limited to lowest order. Consideration may thus be restricted to spatial phase dependence, in angle $\phi = \vec{k}\cdot\vec{r}$ of the complex exponent. The phase field is conserved or Lie-dragged[SCHU 80] under perturbation of deviations $\delta\vec{a}$ & $\delta\vec{z}$:

$$\delta\phi_{\delta\vec{z}}(\vec{a}) \;=\; \frac{2\pi}{\lambda}\,(r_{az} - r_{az_0}) \quad \text{due only to } \delta\vec{z} \tag{A.90}$$

$$r_{az} \;\sim\; r_{a_0 z} - \delta\tilde{a}\cdot\vec{\tilde{r}}_{a_0 z} \quad \text{to reference point } \vec{z}_0$$

$$r_{az_0} \;\sim\; r_{a_0 z} - \delta\tilde{a}\cdot\vec{\tilde{r}}_{a_0 z_0} \quad \text{to perturbed point } \vec{z}$$

$$\vec{\tilde{r}}_{a_0 z} \;\sim\; \vec{\tilde{r}}_{a_0 z_0} + \frac{1}{r_{a_0 z_0}}\,[\delta\vec{z} - (\vec{\tilde{r}}_{a_0 z_0}\cdot\delta\vec{z})\vec{\tilde{r}}_{a_0 z_0}]$$

$$\delta\phi_{\delta\vec{z}}(\vec{a}) \;\sim\; \frac{2\pi}{\lambda r_{a_0 z_0}}\,[(\delta\tilde{a}\cdot\delta\vec{z}) - (\delta\tilde{a}\cdot\vec{\tilde{r}}_{a_0 z_0})(\vec{\tilde{r}}_{a_0 z_0}\cdot\delta\vec{z})]$$

$$\delta\phi_{\delta\vec{z}}(\vec{a}) \sim \frac{2\pi}{\lambda r_{a_0 z_0}}\,\delta\tilde{a}\cdot\delta\vec{z}_{\perp} \quad \text{where} \quad \delta\vec{z}_{\perp} \;=\; (\mathbf{1} - \vec{\tilde{r}}_{a_0 z_0}\otimes\vec{\tilde{r}}_{a_0 z_0})\,\delta\vec{z}$$

Expression of the (asymptotic) auto-correlation follows, by Wiener integration, with asymptotic functional phase shift.[EMER 89] The results may be expressed, in terms of intensity tensor $\mathcal{I}$ and scattering source spatial distribution $p(\vec{a})$, as follows:

$$(\vec{\tilde{k}}\cdot\vec{\tilde{n}})^{-2}\,\langle\, i(\vec{x})\, i(\vec{y})\,\rangle \;\sim\; [\sigma\,(\tfrac{\omega}{c})^2]^2\,\langle\,(\tilde{A}(\vec{x})\cdot\vec{A}^{*}(\vec{x}))(\tilde{A}(\vec{y})\cdot\vec{A}^{*}(\vec{y}))\,\rangle \tag{A.91}$$

$$\sim\; [\sigma\,(\tfrac{\omega}{c})^2]^2\,\langle\,(\int_{\vec{a}} d\tilde{A}(\vec{x})\cdot\int_{\vec{a}} d\vec{A}^{*}(\vec{x}))(\int_{\vec{a}} d\tilde{A}(\vec{y})\cdot\int_{\vec{a}} d\vec{A}^{*}(\vec{y}))\,\rangle$$

$$\sim\; [\sigma\,(\tfrac{\omega}{c})^2]^2\,\langle\,(\int_{\vec{a}} d\tilde{A}\cdot\int_{\vec{a}} d\vec{A}^{*})(\int_{\vec{a}} e^{+i\delta\phi_{\vec{y}-\vec{x}}(\vec{a})} d\tilde{A}\cdot\int_{\vec{a}} e^{-i\delta\phi_{\vec{y}-\vec{x}}(\vec{a})} d\vec{A}^{*})\,\rangle$$

$$\sim\; trace^2(\mathcal{I}) + \rho(\,i(\vec{x}),\, i(\vec{y})\,)\cdot trace(\mathcal{I}\mathcal{I})$$

$$\rho(\,i(\vec{x}),\, i(\vec{y})\,) \;=\; (2\pi)^3\,|\,\mathcal{F}_{\vec{a}\to\vec{s}}\{p(\vec{a})\}(\vec{s})\,|^2 \quad \text{where} \quad \vec{s} = \frac{2\pi}{\lambda r}\,(\mathbf{1} - \vec{\tilde{r}}\otimes\vec{\tilde{r}})\,(\vec{x}-\vec{y})$$

Properties of the Fourier transform, per equation A.26, with source normalization, per equation A.89, suffice for satisfaction of the trivial near and far signal limits in equation A.88. Of further interest is a cross-correlation

expression for a special intensity distribution. In the case of a Gaussian stochastic source, observed source intensity is distributed per equation B.17, as $p(\vec{a}) = G_3(\vec{a}; \vec{\mu}_{sce}, \vec{\vec{\sigma}}^2_{sce})$, from which following cross-correlation results:

$$\rho(i(\vec{x}), i(\vec{y})) \sim \exp(-\frac{1}{2}(\vec{x}-\vec{y})\cdot\sigma^{2\,-1}_{speck}(\frac{1}{2}(\vec{x}+\vec{y}) - \vec{\mu}_{sce}; \sigma^2_{sce}, \lambda)\cdot(\vec{x}-\vec{y})) \quad \text{(A.92)}$$

$$\sigma^{2\,-1}_{speck}(\vec{r}; \sigma^2_{sce}, \lambda) = 2(\frac{2\pi}{\lambda|\vec{r}|})^2(1-\hat{\vec{r}}\otimes\hat{\vec{r}})\sigma^2_{sce}(1-\hat{\vec{r}}\otimes\hat{\vec{r}}) \quad \text{where}$$

$$\hat{\vec{r}} = \frac{\vec{r}}{|\vec{r}|} \quad \text{asymptotically for} \quad \lambda \ll \sqrt{trace(\sigma^2_{sce})} \ll r$$

One notes that variation of function $f(\Omega)$ and conjugate $f^*(\Omega)$, for a scattering micro-source, should indeed be negligible on such a scale. As anticipated, stochastic effects dominate in production of laser speckle. One draws a conclusion in relation to scattering source properties and speckle nature. Stochastic Wiener integration makes no explicit reference to surface qualities, such as composition and roughness. So long as a surface provides approximate stochastic scatter, theory predicts that other source qualities should have little effect on the nature of observed speckle.

For purposes of this thesis, perceived speckle scaling dimension is of interest only as a gross order parameter. The simple statistical goal is that perceived intensity should vary amongst neighbouring pixels, but that measured local averages should be observed as relatively fixed. One may in fact view such a goal as directive for experimental design.

By design, one must satisfy a pair of nontrivial constraints on statistical observations. First, observed pixel dimension should be no more than a few

orders of magnitude larger than the intensity auto-correlation deviation, to ensure that statistical intensity variation be made observable. Second, pixel sample dimension for cross-correlation should be larger than a few times the intensity auto-correlation deviation, in order that cross-correlations be calculated from requisite independent bivariate trials.

By fortunate coincidence, such constraints are mutually compatible. Experimentally, one might choose composite pixels, spanning only a few characteristic auto-correlation lengths. Statistically, intensity variation becomes observable, and the necessarily larger cross-correlation sample dimensions automatically span independent trial regions. Such was the case in work throughout this thesis.

A.6.5 Laser speckle phase interferometry

One would like to generalize the concepts of phase interferometry, expressed in equation A.70, to true interferometer physics. That simple result considered scalar waves in single spatial dimension. With expectations raised by the suggestion of equation A.72, one anticipates that generalization to real waves might be workable. The asymptotically correct interferometer does however involve some sophistication. One must directly evaluate Pearson's cross-correlation for real electromagnetic fields from a stochastic asymptotic plane-wave source pair.

Consider the natural generalization of the simple scalar wave phase offset in equation A.64. Introduction of phase angle offset θ to realistic EM vector

potentials may be similarly denoted by subscript as follows:

$$\vec{A}_\theta = \vec{A}e^{+i\theta} \longrightarrow \vec{A}_\theta(\vec{r},t) = e^{+i\theta}\vec{A}(\vec{r},t) = \vec{A}e^{+i(\vec{k}\cdot\vec{r}-\omega t+\theta)} \quad \text{(A.93)}$$

$$\vec{A}^*_\theta = \vec{A}^*e^{-i\theta} \longrightarrow \vec{A}^*_\theta(\vec{r},t) = e^{-i\theta}\vec{A}^*(\vec{r},t) = \vec{A}^*e^{-i(\vec{k}\cdot\vec{r}-\omega t+\theta)}$$

One sets out to examine the low order moments of real intensity distributions under phase shift. Due to the introduction of incidence and polarization, which remain nontrivial in real interference, direct evaluation for moments of the Poynting vector $\vec{S}_\pm(\vec{r},t)$ of equation A.77 is required. The time-average real Poynting vector for a source pair expresses thus:

$$\begin{aligned}
\langle \vec{S}(\vec{r},t)\rangle_t &= \frac{1}{2}\sigma\langle(\vec{E}\times\vec{B}^* + \vec{E}^*\times\vec{B})\rangle_t \qquad \text{(A.94)}\\
&= \sigma(\frac{\omega}{c})^2[(\vec{A}'\cdot\vec{A}'^*)\hat{k}' + (\vec{A}''\cdot\vec{A}''^*)\hat{k}'']\\
&+ \frac{1}{2}\sigma(\frac{\omega}{c})^2[(\hat{k}'\cdot\vec{A}'')\vec{A}'^* + (\hat{k}'\cdot\vec{A}''^*)\vec{A}']\\
&+ \frac{1}{2}\sigma(\frac{\omega}{c})^2[(\hat{k}''\cdot\vec{A}')\vec{A}''^* + (\hat{k}''\cdot\vec{A}'^*)\vec{A}'']\\
&- \frac{1}{2}\sigma(\frac{\omega}{c})^2[(\vec{A}'\cdot\vec{A}''^*) + (\vec{A}'^*\cdot\vec{A}'')](\hat{k}' + \hat{k}'')
\end{aligned}$$

Evaluation of stochastic (expectation) moments for Poynting vector $\langle\vec{S}(\vec{r},t)\rangle_t$ yields the effect of shifting phase. With interfering *EM* plane waves, $\vec{A}'$ and $\vec{A}''$, phase offset by angles α and β respectively, intensity moments may be calculated directly. The immediate expectation of radiation propagation yields a trivial invariant:

$$\langle\vec{S}(\vec{r},t)\rangle = trace(\mathcal{I}')\hat{k}' + trace(\mathcal{I}'')\hat{k}'' \qquad \text{(A.95)}$$

Expectations for second order moments do however reveal phase shift information. Basic stochastic manipulation, on the Poynting vector of equation

A.94, allows for extraction of the Pearson's cross-correlation, as expressed in equation A.69. The result, with labels altered from equation 2.5, may be expressed in current conventions as follows:

$$\rho(\langle I_{\alpha\beta}\rangle_t, \langle I_{00}\rangle_t) = 1 + \kappa(\mathcal{I}', \mathcal{I}''; \hat{n})\,[\cos(\alpha - \beta) - 1] \qquad \text{(A.96)}$$

$$\kappa(\mathcal{I}', \mathcal{I}''; \hat{n}) = (1 + \frac{\mathcal{T}(\mathcal{I}', \mathcal{I}''; \hat{n})}{\mathcal{S}(\mathcal{I}', \mathcal{I}''; \hat{n})};)^{-1} \quad \text{where}$$

$$\mathcal{S}(\mathcal{I}', \mathcal{I}''; \hat{n}) =$$

$$(\tilde{n}\mathcal{I}'\vec{n})(\tilde{k}'\mathcal{I}''\vec{k}') + 2\cdot(\tilde{n}\mathcal{I}'\vec{k}'')(\tilde{k}'\mathcal{I}''\vec{n}) + (\tilde{k}''\mathcal{I}'\vec{k}'')(\tilde{n}\mathcal{I}''\vec{n}) +$$

$$[\tilde{n}\cdot(\vec{k}' + \vec{k}'')]^2 trace(\mathcal{I}'\mathcal{I}'') - [\tilde{n}\cdot(\vec{k}' + \vec{k}'')][\tilde{n}(\mathcal{I}'\mathcal{I}'' + \mathcal{I}''\mathcal{I}')(\vec{k}' + \vec{k}'')]$$

$$\mathcal{T}(\mathcal{I}', \mathcal{I}''; \hat{n}) =$$

$$\frac{1}{2}(\tilde{n}\cdot\vec{k}')^2\cdot[trace^2(\mathcal{I}') + 3trace(\mathcal{I}'^2)] +$$

$$\frac{1}{2}(\tilde{n}\cdot\vec{k}'')^2\cdot[trace^2(\mathcal{I}'') + 3trace(\mathcal{I}''^2)]$$

Thus derives a solution which respects point intensity of real light, from a coherent stochastic source pair of arbitrary polarization. Having dispensed with the complexities of polarization, the introduction of composite sources and pixels reduces to a triviality. Evaluation reveals the simple net effect of compositeness: observed interferograms appear equivalent to point intensity results taken under optimal conditions.

The general result evaluates by parametrization of point intensity paths. Parametrizations, labelled by X & Y, may be understood to include partition of stochastic micro-source composites in both members of the scattering source pair, and partition measuring the viewed surface. View of the surface

may include spatial pixel dimension and weighting as well for optical blur. One may extrapolate the requisite form of intensity cross-moment from elements of the net partition to the full measure:

$$\langle dI^{\alpha\beta}(X)\, dI^{00}(Y)\rangle \;=\; dS(X,Y)\cos(\alpha-\beta) + dT(X,Y) \quad\longrightarrow \quad \text{(A.97)}$$

$$\begin{aligned}\langle I^{\alpha\beta} I^{00}\rangle &= \langle [\int_X dI^{\alpha\beta}(X)][\int_Y dI^{00}(Y)]\rangle \;=\; \int_{XY}\langle dI^{\alpha\beta}(X)\, dI^{00}(Y)\rangle \\ &= [\int_{XY} dS(X,Y)]\cos(\alpha-\beta) + [\int_{XY} dT(X,Y)] \\ &= S(\{X\},\{Y\})\cos(\alpha-\beta) + T(\{X\},\{Y\})\end{aligned}$$

Assumed here is the asymptotic limit where intensity tensor expectations $\mathcal{I}'$ & $\mathcal{I}''$ and wave vectors $\hat{k}'$ & $\hat{k}''$ remain fixed. Variability should however be allowed in surface normal $\hat{n}$. Angular variability represents surface texture, whereas magnitude variability models local reflectivity. By inference, equation A.97 indeed describes a most general form of composite moment expectation.

One draws a final inference from the present investigation. Composite pixel intensities following equation A.97 satisfy conditions giving the result of equation A.72. The nature of the composite function, $\kappa(\{X\},\{Y\})$, yields as well to simple interpretation. While polarization has introduced measurement angularity and more complex intensity dependence, interferometry following functional equation 1.5 nonetheless respects simple constraints. Function $\kappa(\{X\},\{Y\})$ respects strict bounds and appears as equivalent to an

optimal point intensity result observed under quiescent surface conditions:

$$\text{if} \quad \kappa(X,Y) = (1 + \frac{dT(X,Y)}{dS(X,Y)})^{-1} \quad \text{and} \tag{A.98}$$

$$\kappa(\{X\},\{Y\}) = (1 + \frac{T(\{X\},\{Y\})}{S(\{X\},\{Y\})})^{-1} \quad \text{then}$$

$$m \leq \frac{dT(X,Y)}{dS(X,Y)} \leq M \longrightarrow m \leq \frac{T(\{X\},\{Y\})}{S(\{X\},\{Y\})} \leq M$$

$$k \leq \kappa(X,Y) \leq K \longrightarrow k \leq \kappa(\{X\},\{Y\}) \leq K$$

Composite intensity measurements inherently mitigate local effects such as surface variability in texture and reflectivity. Variability in function $\kappa(\{X\},\{Y\})$ may be damped, without alteration in observed interference amplitude. Macroscopic pixelation and optical blur show precisely such an effect. So long as speckle variation may be observed as significant in video images, one concludes that compositeness should induce not deterioration but rather enhancement in quality of observed interferograms.

Appendix B

Likelihoods

B.1 The Neyman-Pearson lemma

The investigative goal of the study of likelihoods, converse to the study of probability, is inference of which experiment was conducted, based purely on observation of results in sets of random experiments. While the aim of such study should seem familiar to the experimentalist, direct examination of the theory of likelihood does raise some peculiar concepts.[FREU 62]

One may illustrate the likelihood concept by posing the following thought experiment. Suppose that one was given a coin and the associated task of conducting an experiment of 10 tosses. Suppose further, that the experiment resulted in 8 heads and 2 tails. Now, one may ask, what is the most probable result of the next toss? If one assumes a balanced coin, heads and tails outcomes retain equal probabilities. The preceding experimental evidence however suggests otherwise. The experimental coin might likely be

unbalanced. The sample in fact suggests the intuitive (maximum likelihood) prediction, that the coin should come up heads with 80% and tails with 20% probability respectively. One is left asking why such a qualitative description should seem so reasonable.

One may in fact quantify a general concept of likelihood which reflects the qualitative view. The *Neyman-Pearson lemma* amounts to demonstration of this actuality. A few preliminary refinements are required though. One requires strict definition for intuitive concepts regarding statistical tests. From rigid foundations, one may quite elegantly demonstrate the general truth of the intuitive assessment.

Suppose one has a pair of possible hypotheses A & B as to the nature of an experiment. Assume that proposition A amounts to the practical fall-back assumption, often referred to as the *null hypothesis*. The *alternative hypothesis*, comprised by B corresponds to recognition of membership in a distinctly defined subclass on some clear criterion. One would like to find an optimal decision criterion which experimentally separates the populations.

One recognizes an obvious classification based on observation. Given an observation x one may calculate the probabilities of x supposing hypothesis A holds and supposing hypothesis B holds respectively as $L_A(x)$ & $L_B(x)$. One acknowledges the conditional nature of hypothetical probabilities by coining the term *likelihood* for such distributions. One might guess that the higher value of the likelihood pair indicates the better hypothesis.

One acquires clear insight into classification through examination of *conditional probabilities.* For observable events u & v, one defines by $P(u|v) = P(u\&v)/P(v)$ the conditional probability for event u given that one has already observed event v. One notes that sensibly $P(v) \neq 0$, since after all, v has been observed. One may derive, given observed event x, (with $P(x) \neq 0$,) the following ratio of conditional probabilities:

$$\frac{P(A|x)}{P(B|x)} = \frac{P(A\&x)}{P(B\&x)} = \frac{P(A)}{P(B)} \cdot \frac{P(x|A)}{P(x|B)} \tag{B.1}$$

One intuitively identifies $P(x|A) = L_A(x)$ & $P(x|B) = L_B(x)$. The global probability ratio $P(A)/P(B)$ derives from experimental conditions. One may indeed assign, for the event pair A & B, a relative likelihood classification based on experimental observation.

One should like to investigate further the nature of classification. The result of equation B.1 suggests a test criterion and motivation for the study of optimal hypothesis tests. One sensibly defines the assessment of best hypothesis from equation B.1 by $H(x; \{L_A()\}, \{L_B()\}, k)$ on the following functional criterion:

$$H(x; \{L_A()\}, \{L_B()\}, k) = \begin{cases} A & \tau(x; \{L_A()\}, \{L_B()\}) > k \\ B & \tau(x; \{L_A()\}, \{L_B()\}) < k \\ \text{either} & \tau(x; \{L_A()\}, \{L_B()\}) = k \end{cases}$$

$$\tau(x; \{L_A()\}, \{L_B()\}) = \frac{L_A(x)}{L_B(x)} \tag{B.2}$$

Parameter k may be varied to bias preference toward either hypothesis A or B. Choice of k affects relative likelihoods of misclassification by skewing questionable hypothesis comparison ratios preferentially. The constant

may also be identified with global probability ratio $P(A)/P(B)$ appearing in equation B.1.

One may further precise the test definition by writing the test in terms of set R and conjugate R', referred to as *critical regions*, such that:

$$H(x;\{L_A()\},\{L_B()\},C,C') = \begin{cases} A & x \in R \\ B & x \in R' \end{cases} \quad \text{where} \qquad \text{(B.3)}$$

$$R \supseteq \{x : \tau(x;\{L_A()\},\{L_B()\}) > k\}$$

$$R' \supseteq \{x : \tau(x;\{L_A()\},\{L_B()\}) < k\}$$

One may observe that more general functions $\tau(x;\{L_A()\},\{L_B()\})$ may be tested against constants k, and that one may thus define more general critical regions. Conversely, one may view a general test as constructed by choice of its critical regions. The question presents itself, whether other test criteria exist, which show optimal performance (in some sense) relative to the *ratio of likelihoods* test.

The most obvious measure of test performance relates to the amount of classification error. The traditional definition of error types is as follows:

type I: rejection of the null hypothesis when it is true

type II: acceptance of the null hypothesis when it is false

Viewing error classification in terms of error probability allows for a balanced treatment regarding the pair of hypotheses. Type I errors erroneously reject null hypothesis A, comprising measure α under distribution $L_A(x)$;

Type II errors erroneously reject alternative hypothesis B, comprising measure β under distribution $L_B(x)$. One identifies sets where a test fails by the respective errors with precisely the critical regions of that test. One often refers to these by measure or size. One infers from this balanced view that a sensible definition of optimality amounts to minimization of either error, with the other held fixed. One finds that choice as to which error is minimized and which is held fixed has no effect on the result.

The *ratio of likelihoods* test criterion of equation B.2 shows itself as optimal by such error minimization. The result may be demonstrated directly as follows. Suppose that *ratio of likelihoods* test (for some k) has critical region R of of type I error size α; more simply, R defines the set where hypothesis A fails on the *ratio of likelihoods* test. Consider also any other test with respective critical region S of identical size α. Symbollically, one may write the fixed error term as:

$$\int_{r \in R} dr\, L_A(r) \;=\; \alpha \;=\; \int_{s \in S} ds\, L_A(s) \tag{B.4}$$

By cancelling integrations over shared region $R \cap S$ one finds:

$$\int_{t \in R \cap S'} dt\, L_A(t) \;=\; \int_{t \in R' \cap S} dt\, L_A(t) \tag{B.5}$$

One next observes from equation B.2, that for any positive choice of k, the following pair of identities hold on the respective regions:

$$\begin{aligned} x \in R \cap S' \;&\longrightarrow\; L_B(x) \;\geq\; \frac{1}{k} L_A(x) \\ x \in R' \cap S \;&\longrightarrow\; \frac{1}{k} L_A(x) \;\geq\; L_B(x) \end{aligned} \tag{B.6}$$

As a result, one derives the following meaningful inequality:

$$\int_{t\in R\cap S'} \mathbf{d}t\, L_B(t) \geq \frac{1}{k}\int_{t\in R\cap S'} \mathbf{d}t\, L_A(t) = \frac{1}{k}\int_{t\in R'\cap S} \mathbf{d}t\, L_A(t) \geq \int_{t\in R'\cap S} \mathbf{d}t\, L_B(t) \qquad \text{(B.7)}$$

One may apply precisely this inequality to generate a result regarding the likelihoods of avoiding type II errors:

$$\begin{aligned} 1-\beta_R &= \int_{t\in R} \mathbf{d}t\, L_B(t) = \int_{t\in R\cap S} \mathbf{d}t\, L_B(t) + \int_{t\in R\cap S'} \mathbf{d}t\, L_B(t) \\ &\geq \int_{t\in R\cap S} \mathbf{d}t\, L_B(t) + \int_{t\in R'\cap S} \mathbf{d}t\, L_B(t) = \int_{t\in S} \mathbf{d}t\, L_B(t) = 1-\beta_S \end{aligned} \qquad \text{(B.8)}$$

One may conclude the demonstration by noting that $\beta_R \leq \beta_S$: likelihood of type II error under the *ratio of likelihoods* test gives optimal performance.

One may expand the result by viewing both the above demonstration and its converse equivalent. By interchanging set labels one may swap critical regions for null and alternate hypotheses, showing optimality as well in the converse view, where error measure β (type II error) is held fixed and α (type I error) is optimized. One labels a test satisfying such characteristics of optimality as having *most powerful critical regions.*

In conclusion of this introduction to the concept of likelihood, one notes that intuitive notions of qualitative likelihood may be sensibly quantified by adherence to accepted definitions and conventions. Concepts which seem awkward and ill-defined in a qualitative view become well grounded in the peculiar rigorous view of the preceding demonstration.

B.2 Maximum likelihood methods

Maximum likelihood methods encompass procedures where functional optimizations are drawn from analysis based on the *Neyman-Pearson Lemma* of the previous section. As a special case of equation B.1, one may compare experimental likelihoods for similar parametrized distribution models. One assumes equal apriori global probabilities $P(A') = P(A'')$ for test distributions parametrized by $\vec{\alpha}'$ & $\vec{\alpha}''$ respectively. Given experimentally observed data sample $\vec{x}$ one may express conditional probability ratios as follows:

$$\frac{P(A'|\vec{x})}{P(A''|\vec{x})} = \frac{P(\vec{x}|A')}{P(\vec{x}|A'')} = \frac{p(\vec{x};\vec{\alpha}')}{p(\vec{x};\vec{\alpha}'')} \tag{B.9}$$

By testing parametrizations pairwise according to equation B.9, a unique and consistent optimality ordering arises. Moreover, as parameters $\vec{\alpha}$ vary over their allowed domain, the function $p(\vec{x};\vec{\alpha})$ may be maximized. Continuity of $p(\vec{x};\vec{\alpha})$ over a compact domain of test parameters $\vec{\alpha}$ suffices to ensure that a maximum is actually attained. One thereby chooses, from amongst a class of distributions parametrized by $\vec{\alpha}$, the optimal member distribution given experimental result $\vec{x}$. Functional classification error assumes optimimality for such a *maximum likelihood* result.

Practical applications of *maximum likelihood* often involve data sets with iterated independent trials. In the case of such an independent iteration, general optimization obtains from expression of the likelihood of the complete

experimental result, given a choice of parameters:

$$P(\vec{x};\vec{\alpha}) = \prod_{n=1}^{n=N} L(x_n;\vec{\alpha}) \tag{B.10}$$

$$\vec{0} = \frac{\partial}{\partial\vec{\alpha}} \ln P(\vec{x};\vec{\alpha})$$

One solves the latter system of equations to recover *maximum likelihood* $\vec{\alpha}$ parameters.[MATH 70]

As example for the method, assume independent data of Gaussian distribution, per equation A.12. For a fixed data set of repeated trial measurements, x_n with $n \in \{1,\ldots,N\}$, one may vary Gaussian parameters μ & σ to find optimal values:

$$P(\vec{x};\mu,\sigma) = \prod_{n=1}^{n=N} g(x_n;\mu,\sigma) \tag{B.11}$$

One optimizes $P(\vec{x};\mu,\sigma)$ by solving following systems of equations:

$$0 = \frac{\partial}{\partial\mu} \ln P(\vec{x};\mu,\sigma) \longrightarrow \mu = \frac{1}{N}\sum_{n=1}^{n=N} x_n \tag{B.12}$$

$$0 = \frac{\partial}{\partial\sigma} \ln P(\vec{x};\mu,\sigma) \longrightarrow \sigma^2 = \frac{1}{N}\sum_{n=1}^{n=N} (x_n-\mu)^2 \tag{B.13}$$

The *maximum likelihood mean* of equation B.12 equates identically with the unbiased estimate of equation A.9, while the *maximum likelihood variance* of equation B.13 scales by a factor of $\frac{N-1}{N}$ relative to the unbiased estimate of equation A.10. Introduction of the corresponding bias gives a better variance estimate for the particular data set in question. Statisticians classify such optimizations as *point estimation.*

More general applications abound. Any independent statistical measurements may express products in the form of equation B.10, with parameter variation optimizing likelihood. Further examples of *maximum likelihood methods* appear in the following appendices.

B.3 Pearson's cross-correlation

With the concept of a bivariate probability distribution, $P(x, y)$, one encounters mixed moments of the form $\langle x^m y^n \rangle$, lending new and robust potentials to the study of expectations. New non-negative quantities are readily constructed: given general $a\&b \in \mathcal{R}$, one finds the following example:

$$0 \leq \langle [a(x-\bar{x}) + b(y-\bar{y})]^2 \rangle = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 \\ \sigma_{yx}^2 & \sigma_{yy}^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{(B.14)}$$

The *correlation tensor* must satisfy the non-negative semi-definite condition. Most simply stated, eigenvalues of the tensor must be non-negative. Equivalently, $0 \leq \sigma_{xx}^2$ & $0 \leq \sigma_{yy}^2$, and the tensor's determinant must be non-negative as well:

$$0 \leq \det \sigma^2 = \sigma_{xx}^2 \cdot \sigma_{yy}^2 - \sigma_{xy}^2 \cdot \sigma_{yx}^2 \quad \text{(B.15)}$$

One thus finds for cases with nontrivial statistics, where ($\sigma_{xx}^2 > 0$ & $\sigma_{yy}^2 > 0$), the following result:

$$\rho = \frac{\sigma_{xy}^2}{\sqrt{\sigma_{xx}^2 \, \sigma_{yy}^2}} \quad \longrightarrow \quad -1 \leq \rho \leq +1 \quad \text{(B.16)}$$

Clearly inferred, *Pearson's cross-correlation* ρ thus defined, bears information content. One may in fact draw conclusions based solely on this

statistic. If random variables x and y are independent of one another, then $\rho(x,y) = 0$. Thus non-vanishing ρ contradicts *pairwise independence* of x & y. At the other extreme, $\rho(x,y) = \pm 1$ corresponds to an exact linear relationship: $ax + by = c$. The relationship holds exactly, except possibly on a set of zero (probability) measure. Experimentally, moreover, deviations from perfection may be measured as caused by statistical error.

One might next ask how best to estimate cross-correlations from experimental samples. Derivation of a *maximum likelihood estimate* for ρ follows the prescription of equation B.10. Solution here makes use of the general multidimensional Gaussian having vector mean $\vec{\mu}$ and positive-definite covariance tensor $\vec{\vec{\sigma}}^2$:[PATH 72]

$$G_N(\vec{x};\vec{\mu},\vec{\vec{\sigma}}^2) = \frac{1}{\sqrt{(2\pi)^N \cdot \det \sigma^2}} \exp[-\frac{1}{2}(\tilde{x}-\tilde{\mu})\,\sigma^{2\,-1}(\vec{x}-\vec{\mu})] \qquad \text{(B.17)}$$

An experimental likelihood, over independent multidimensional Gaussian trials, expresses as a product of the following form:

$$P(\vec{\vec{x}};\vec{\mu},\vec{\vec{\sigma}}^2) = \prod_{n=1}^{n=N} G_K(\vec{x}_n;\vec{\mu},\vec{\vec{\sigma}}^2) \qquad \text{(B.18)}$$

Optimization of $P(\vec{\vec{x}};\vec{\mu},\vec{\vec{\sigma}}^2)$ follows by solving systems of equations:[FREU 62]

$$0 = \frac{\partial}{\partial\vec{\mu}} \ln P(\vec{\vec{x}};\vec{\mu},\vec{\vec{\sigma}}^2) \longrightarrow \vec{\mu} = \frac{1}{N}\sum_{n=1}^{n=N} \vec{x}_n \qquad \text{(B.19)}$$

$$0 = \frac{\partial}{\partial\vec{\vec{\sigma}}^2} \ln P(\vec{\vec{x}};\vec{\mu},\vec{\vec{\sigma}}^2) \longrightarrow \vec{\vec{\sigma}}^2 = \frac{1}{N}\sum_{n=1}^{n=N} (\vec{x_n}-\vec{\mu}) \otimes (\vec{x_n}-\vec{\mu}) \qquad \text{(B.20)}$$

Results B.19 and B.20 mirror the respective simple single dimensional counterparts of equations B.12 and B.13.

In the example of the 2-dimensional case, with $x_1 = x$ &$x_2 = y$, parameters transform as $(\sigma^2_{xx}, \sigma^2_{xy}, \sigma^2_{yy}) \longrightarrow (\sigma^2_{xx}, \rho, \sigma^2_{yy})$ without altering any likelihood values. One thus finds *maximum likelihood* cross-correlation:

$$\rho(x,y) = \frac{N\sum xy - \sum x \sum y}{\sqrt{(N\sum xx - \sum x \sum x)(N\sum yy - \sum y \sum y)}} \tag{B.21}$$

The *maximum likelihood estimate* of expression B.21 (with indices suppressed for clarity) has been used throughout this thesis.

Multivariate distributions allow as well for calculation of further quantities. Linear functions of random variables generally follow:[MATH 70]

$$\begin{aligned} z = \quad & z_0 + \sum_k \frac{\partial z}{\partial s^k} \cdot s_k \quad = \quad z_0 + \frac{\partial z}{\partial \vec{s}} \cdot \vec{s} \quad \longrightarrow \\ \bar{z} = \quad & z_0 + \sum_k \frac{\partial z}{\partial s^k} \cdot \bar{s}_k \quad = \quad z_0 + \frac{\partial z}{\partial \vec{s}} \cdot \bar{\vec{s}} \\ \sigma^2_{zz} = \quad & \sum_{mn} \frac{\partial z}{\partial s^m} \cdot \sigma^2_{mn} \cdot \frac{\partial z}{\partial s^n} \quad = \quad \frac{\partial z}{\partial \vec{s}} \cdot \boldsymbol{\sigma}^2 \cdot \frac{\partial z}{\partial \tilde{s}} \end{aligned} \tag{B.22}$$

Here $\boldsymbol{\sigma}^2$ gives the matrix equivalent of the covariance tensor for purposes of the calculation. The result may be extended to nonlinear functions, with negligible second order contributions, giving *linearized variance* estimates.

As an example of a linearized variance (error) estimate, consider the sample cross-correlation of equation B.21 as a function of data vectors $\vec{x}$ & $\vec{y}$. Assuming statistically independent measurement pairs, one has:

$$\begin{aligned} \sigma^2_{x_m x_n} &= \delta_{mn} \cdot \sigma^2_{xx} \\ \sigma^2_{x_m y_n} &= \delta_{mn} \cdot \sigma^2_{xy} \\ \sigma^2_{y_m y_n} &= \delta_{mn} \cdot \sigma^2_{yy} \end{aligned} \tag{B.23}$$

Application of linearized variance, per equation B.22, to the sample cross-correlation equation B.21, respecting the covariances of equation B.23, yields a surprisingly simple result:

$$\sigma^2_{\rho\rho} \;=\; \frac{2}{N}\,(1-\rho^2)^2 \qquad \text{(B.24)}$$

Variance in the cross-correlation may be expressed, to lowest order, as a function of the value on itself. Effects from data distribution cancel in the linear calculation.

Equation B.24 raises a measurable consequence. One may investigate the assumption of independent measurements expressed in equation B.23 by measuring real sample ρ and $\sigma^2_{\rho\rho}$. Effective and actual sample populations (N from equation B.24) should correspond. Detailed theoretical and experimental examinations appear respectively in subsections 2.1.4 and 4.2.2.

B.4 General linear regression

The general least squares problem may be expressed as an application of *maximum likelihood methods*. Optimization for a linear regression reduces to choice of a (positive definite) quadratic. The quadratic χ- statistic should be chosen to vanish identically, when theory and experiment match. The remainder of the optimization follows as trivial algebra.[MATH 70][PRES 92]

Suppose one wishes to test a theoretical prediction of the following form:

$$\phi(\vec{z}) \;=\; \tilde{a}\cdot\vec{f}(\vec{z}) \qquad \text{(B.25)}$$

Here, one assumes (well defined) measurement function $\phi(\vec{z})$, and order n set of (well defined) basis functions $f_k(\vec{z})$, being linearly independent on the (as yet undefined) experimental data set. One wishes to measure the n corresponding (dual) coefficients a_k. One should, therefore, choose (positive definite) quadratic form χ as follows:

$$\chi_{\text{exp}} \;=\; \langle\, [\phi(\vec{z}) - \vec{a}\cdot\vec{f}(\vec{z})]^2 \,\rangle_{\text{exp}} \tag{B.26}$$

The suggested form vanishes with perfect theoretical data, and assumes a unique minimum stationary point for noisy data. As will be made clear in the following exposition, that stationary point gives the best fit.

Experimentally, solution requires $N \geq n$ independent trials. One assumes the experimental data set to be nonsingular, in a sense which will become clear below. In anticipation of the following discussion, one has defined a new *experimental* expectation $\langle \ldots \rangle_{\text{exp}}$ with following normalization:

$$\begin{aligned}
\langle\, \Phi(\vec{z};k) \,\rangle_{\text{exp}} \;&=\; \sum_{k=1}^{k=N} \frac{1}{\sigma_{kk}^2}\, \Phi(\vec{z}_k;k) \quad \text{where} \\
\langle\, \Delta_m \cdot \Delta_n \,\rangle \;&=\; \delta_{mn}\, \sigma_{mn}^2 \quad \text{statistical expectation with} \\
\Delta_k(\vec{z};\vec{a}_{\text{theory}}) \;&=\; \phi(\vec{z_k}) - \vec{a}_{\text{theory}}\cdot\vec{f}(\vec{z_k}) \quad k \in \{1,\ldots,N\}
\end{aligned} \tag{B.27}$$

Optimization reduces to solution for the (unique) stationary point of the quadratic form in equation B.26. Standard solution may be expressed by:

$$\vec{0} \;=\; \frac{1}{2}\cdot\frac{\partial}{\partial\vec{a}}\,\chi_{\text{exp}} \;=\; \sigma^{2\,-1}\vec{a}_{\text{fit}} - \vec{v} \quad \text{where} \tag{B.28}$$

$$\sigma_{ij}^{2-1} \;=\; \langle\, f_i(\vec{z})\cdot f_j(\vec{z}) \,\rangle_{\text{exp}} \quad \text{and} \quad v_i \;=\; \langle\, \phi(\vec{z}) f_i(\vec{z}) \,\rangle_{\text{exp}}$$

$$\vec{v} = \sigma^{2\,-1}\vec{a}_{\rm fit} \quad \& \quad \exists\sigma^2\colon\ \sigma^2 = (\sigma^{2\,-1})^{-1} \longrightarrow \vec{a}_{\rm fit} = \sigma^2\,\vec{v}$$

One infers, by way of solution, that (non-negative semi-definite) symmetric form $\sigma^{2\,-1}$ (ie. $\sigma^{2-1}_{ij} = \sigma^{2-1}_{ji}$) should be (positive-definite) invertible. Such requirement may in some sense be circumvented by the more general *singular value decomposition* presented in section 3.4. Conventional single valued solution requires that data be non-singular in the sense that $\sigma^{2\,-1}$ be matrix-invertible.

Provided a solution exists, as outlined in equation B.28, equation B.26 may be rewritten as follows:

$$\begin{aligned}
\chi_{\rm fit}(\{\vec{z}\};\vec{a}) &= \chi^0_{\rm fit}(\{\vec{z}\}) + (\tilde{a}-\tilde{a}_{\rm fit})\,\sigma^{2\,-1}(\{\vec{z}\})\,(\vec{a}-\vec{a}_{\rm fit}) \quad \text{where} \quad \text{(B.29)}\\
\chi^0_{\rm fit}(\{\vec{z}\}) &= \langle\,\phi^2(\vec{z})\,\rangle_{\rm exp} - \tilde{a}_{\rm fit}(\{\vec{z}\})\cdot\vec{v}(\{\vec{z}\}) \quad \text{independent of } \vec{a}\\
\vec{a}_{\rm fit}(\{\vec{z}\}) &= \sigma^2(\{\vec{z}\})\,\vec{v}(\{\vec{z}\}) \quad \text{linear regression fit}
\end{aligned}$$

One recognizes the definition of $\chi^0_{\rm fit}$ as a multi-dimensional generalization to the variance expression of equation A.6.

Application of the *maximum likelihood principle* reduces to simple inference from the preceding. If one presupposes Gaussian distribution for variables Δ_k of equation B.27, one infers the following likelihood product:

$$\begin{aligned}
P(\vec{\Delta};\vec{a}) &= \prod_{k=1}^{k=N} g(\,\Delta_k(\vec{z}_k;\vec{a});\,\mu=0,\,\sigma=\sqrt{\sigma^2_{kk}}\,) \qquad \text{(B.30)}\\
&= \frac{1}{\sqrt{(2\pi)^N\prod_k\sigma^2_{kk}}}\,\exp(\,-\frac{1}{2}\,\langle\,\Delta^2_k(\vec{z}_k;\vec{a})\,\rangle_{\rm exp}\,)
\end{aligned}$$

One observes that the exponent matches identically with the experimental

expectation term of equation B.26. Thus maximum likelihood analysis, via equation B.10 and respecting equation B.30, gives the same solution for $\vec{a}_{\rm fit}$ as expressed by the respective stationary point solution of equation B.28. In this instance, maximum likelihood and least-squares solutions are perfectly interchangeable.

From a slightly more sophisticated vantage, one would like to presuppose for the experiment, multidimensional Gaussian likelihood of another format:

$$P(\vec{\delta}_{\rm noise}, \vec{a}_{\rm fit}; \vec{a}_{\rm theory}, \sigma_{\delta\delta}^{2\ \rm noise}, \sigma_{aa}^{2\ \rm theory}) \;=\; \frac{\exp(-\frac{1}{2}\,\chi_{\rm theory})}{\sqrt{(2\pi)^N \det\sigma_{\delta\delta}^2 \det\sigma_{aa}^2}} \qquad \text{(B.31)}$$

$$\chi_{\rm theory} = \chi^0_{\rm theory} + (\tilde{a}_{\rm fit} - \tilde{a}_{\rm theory})\,\sigma_{aa}^{2\ -1}\,(\vec{a}_{\rm fit} - \vec{a}_{\rm theory}) \quad \text{where} \quad \chi^0_{\rm theory} = \tilde{\delta}\,\sigma_{\delta\delta}^{2\ -1}\vec{\delta}$$

If such an identity holds, one may view in separation, noise modes $\vec{\delta}_{\rm noise}$ as independent of the random variables $\vec{a}_{\rm fit}$. This view accords well with intuitive concepts of measurement in the presence of noise.

Generally, likelihoods expressed in equations B.30 and B.31 are only approximately compatible. One could investigate the relationship between the likelihood expressions by attempting to construct the requisite transformation from one into the other. Most sensibly, here, one should express both with respect to distribution of the underlying $\vec{z}$. Such investigation, though, extends well beyond the needs of this thesis. Relying upon such a qualitative view, however, one may exploit the natural expression of equation B.31, to draw some meaningful inferences.

Comparing the experimental χ of equation B.29 with the model $\chi_{\rm theory}$ of

equation B.31, one recovers *maximum likelihood* estimate $\vec{a}_{\text{theory}} \sim \vec{a}_{\text{fit}}(\{\vec{z}\})$. Comparing further coefficients one infers $\sigma^2_{aa} \sim \sigma^2(\{\vec{z}\})$ as covariance (error) tensor. One moreover postulates following measure for degrees of freedom:

$$\begin{aligned} n &= \langle (\tilde{a}_{\text{fit}} - \tilde{a}_{\text{theory}})\, \sigma^{2\,-1}_{aa} (\vec{a}_{\text{fit}} - \vec{a}_{\text{theory}}) \rangle \\ N - n &= \langle \tilde{\delta}\, \sigma^{2\,-1}_{\delta\delta} \vec{\delta} \rangle = \langle \chi^0_{\text{theory}} \rangle \sim \langle \chi^0_{\text{fit}} \rangle \sim \chi^0_{\text{fit}}(\{\vec{z}\}) \end{aligned} \tag{B.32}$$

From such expressions one may estimate, even without apriori knowledge of error magnitudes, the elements of the covariance (error) tensor σ^2_{aa}. More precisely, if one assumes weights to be known only up to a constant scaling factor, one may eliminate that factor as follows:

$$\sigma^2_{aa} \sim \frac{\langle \chi^0_{\text{fit}} \rangle}{N-n} \langle \vec{f}(\vec{z}) \otimes \tilde{f}(\vec{z}) \rangle^{-1}_{\text{exp}} \sim \frac{\chi^0_{\text{fit}}(\{\vec{z}\})}{N-n} \sigma^2(\{\vec{z}\}) \tag{B.33}$$

Such assignment of experimental values to fill unknown expectation values amounts to *point estimation*. One ordinarily assumes experiments with data size nearing the limit $N \gg n$. In such cases, distributions are sharply peaked, and point estimates most often suffice for purposes of error analysis.

The foregoing presentation has demonstrated linear regression as a pure application of *maximum likelihood*. Other approaches do exist, however, which apply conventional algebraic techniques. Any solution for best fit vector $\vec{a}_{\text{fit}}$ should naturally return identically the same equation B.28 result.

The identical covariance (error) tensor σ^2_{aa} often derives simply as well from a linearized calculation.[MATH 70] Calculation here relies upon the

following variational approximation:

$$\begin{aligned} \vec{a}(\{\vec{z}\}) &= \sigma^2(\{\vec{z}\})\,\vec{v}(\{\vec{z}\}) \quad \longrightarrow \\ \delta\vec{a} &= \sigma^2\delta\vec{v} - \sigma^2[\delta(\sigma^{2\,-1})]\,\sigma^2\vec{v} \;\sim\; \sigma^2\delta\vec{v} \\ \delta\vec{v} &= \langle\, \vec{f}\,\delta\phi(\vec{z}) + \phi\,\delta\vec{f}(\vec{z})\,\rangle_{\rm exp} \;\sim\; \langle\, \vec{f}(\vec{z})\,\delta\phi(\vec{z})\,\rangle_{\rm exp} \end{aligned} \tag{B.34}$$

For validity of such approximation, errors in the dependent (measured) quantity should dominate. More precisely, functions $\phi(\vec{z})$ & $\vec{f}(\vec{z})$ (per equation B.25) should satisfy $\sigma^2_{\phi\phi}(\vec{z}) \gg |\sigma^2_{ff}(\vec{z})|$ for each measurement. Linearized covariance estimation may then be expressed in matrix equivalent form:

$$\begin{aligned} \sigma^2_{aa} &\sim \langle\, \delta\vec{a}\otimes\delta\vec{a}\,\rangle \sim \langle\, (\delta\vec{v}\,\sigma^2)\otimes(\sigma^2\delta\vec{v})\,\rangle \sim \langle\, \sigma^2(\delta\vec{v}\otimes\delta\vec{v})\,\sigma^2\rangle \\ &\sim \langle\, \sigma^2\langle\vec{f}(\vec{z})\otimes\vec{f}(\vec{z})\,\rangle_{\rm exp}\,\sigma^2\rangle \sim \langle\, \sigma^2\,\sigma^{2\,-1}\sigma^2\rangle \\ &\sim \langle\, \sigma^2\rangle \sim \sigma^2(\{\vec{z}\}) \end{aligned} \tag{B.35}$$

A sensible linearized calculation does indeed confirm that covariance (error) tensor follows $\sigma^2_{aa} = \sigma^2$, in good agreement with equation B.31 assertion.

One concludes that relatively simple conditions allow for recovery of best fit vector $\vec{a}_{\rm fit}$ and corresponding covariance (error) tensor σ^2_{aa}. Equations B.28 and B.33 demonstrate how sensible measurement precepts suffice for identification of respective experimental quantities. Generality of the solution extends, however, with concepts from this presentation of *linear regression* analysis seeing routine use in a vast diversity of applications.

Appendix C

Computational methods

C.1 Least squares stability analysis

In applications involving computation, one often encounters problems at the limits of machine precision.[PRES 92] A standard reaction on encountering precision related issues is to move from single precision to double precision floating point math in performance of the critical task. This approach often fails to address the causative factors leading to precision loss. The true source of such problems often rests in the algorithm performing the critical calculation, rather than in the precision of stored data.

One may illustrate the point with a rather simple example. Consider the standard linear regression. One may write the χ- statistic and solution in the notation of Appendix B.4 as follows:

$$\chi = \langle (y - a - bx)^2 \rangle \quad \text{and} \quad \Delta = \langle x^2 \rangle - \langle x \rangle^2 \quad \longrightarrow \qquad \text{(C.1)}$$

$$\begin{bmatrix} \langle y \rangle \\ \langle xy \rangle \end{bmatrix} = \begin{bmatrix} \langle 1 \rangle & \langle x \rangle \\ \langle x \rangle & \langle x^2 \rangle \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} +\langle x^2 \rangle & -\langle x \rangle \\ -\langle x \rangle & +\langle 1 \rangle \end{bmatrix} \begin{bmatrix} \langle y \rangle \\ \langle xy \rangle \end{bmatrix}$$

One immediately notes a pair of opportunities for precision loss. First, calculation of sums for tensor and vector elements might add components of disparate magnitudes or even worse cancel large values by subtraction. Second, calculation of the determinant certainly cancels a value of large magnitude by subtraction. In both cases floating point information is lost.

Precision losses of the types illustrated above are needless. One quickly demonstrates the point by recasting coordinates. About the centroid, with $\bar{x} = \langle x \rangle$ and $\bar{y} = \langle y \rangle$, the identical problem may be expressed as follows:

$$\chi = \langle (\delta y - a' - b'\delta x)^2 \rangle \quad \text{where} \quad \delta x = x - \bar{x} \text{ and } \delta y = y - \bar{y} \quad \longrightarrow \quad \text{(C.2)}$$

$$\begin{bmatrix} 0 \\ \sigma_{xy}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sigma_{xx}^2 \end{bmatrix} \begin{bmatrix} a' \\ b' \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sigma_{xx}^{2\,-1} \end{bmatrix} \begin{bmatrix} 0 \\ \sigma_{xy}^2 \end{bmatrix}$$

$$a' = 0 \text{ and } b' = \frac{\sigma_{xy}^2}{\sigma_{xx}^2} \quad \longrightarrow \quad a = \bar{y} - b\bar{x} \text{ and } b = b'$$

Inspection of equation C.2 shows that the parameter and coordinate transformation reduces linear regression to a mathematical triviality. The new prescription also eliminates both sources of possible precision loss present in the common regression of equation C.1. The method works by first minimizing magnitudes in moment calculations and second trivializing the fit determinant.

Similar methods may be employed in more complicated multiple regression analysis, and even in cases where singular value decompositions are required. One finds in fact, that as fit dimension increases, reliance on algorithm stability becomes an increasingly critical issue.

Application of a transformation along the lines of equation C.2 may be simply accomplished if one of the fit functions gives a constant offset. In that case, one may write the χ- statistic for the general multiple regression as follows:

$$\chi_{\rm exp} = \langle [\phi(\vec{z}) - \tilde{a} \cdot \vec{f}(\vec{z})]^2 \rangle_{\rm exp} = \langle [\delta\phi(\vec{z}) - \tilde{a}' \cdot \delta\vec{f}(\vec{z})]^2 \rangle_{\rm exp} \quad \text{where} \quad \text{(C.3)}$$

$$f_0(\vec{z}) = 1;\ \bar{\phi} = \frac{\langle \phi(\vec{z}) \rangle_{\rm exp}}{\langle 1 \rangle_{\rm exp}};\ \delta\phi(\vec{z}) = \phi(\vec{z}) - \bar{\phi};\ \bar{\vec{f}} = \frac{\langle \vec{f}(\vec{z}) \rangle_{\rm exp}}{\langle 1 \rangle_{\rm exp}};\ \delta\vec{f}(\vec{z}) = \vec{f}(\vec{z}) - \bar{\vec{f}}$$

$$\longrightarrow \quad a_i = a_i' \ \forall i \neq 0 \quad \text{and} \quad a_0 = \bar{\phi} - \sum_{i \neq 0} \bar{f}_i \quad \text{where} \quad a_0' = 0$$

The transformation in equation C.3 has accomplished as well conversion of the curvature tensor into a true covariance matrix. Diagonal elements are thus minimized, and other elements are bounded by the non-negagive semi-definite nature of the covariance tensor:

$$\langle \delta\vec{f}(\vec{z}) \otimes \delta\tilde{f}(\vec{z}) \rangle_{\rm exp} = \langle \vec{f}(\vec{z}) \otimes \tilde{f}(\vec{z}) \rangle_{\rm exp} - \bar{\vec{f}} \otimes \bar{\tilde{f}} \langle 1 \rangle_{\rm exp} \qquad \text{(C.4)}$$

The subtracted tensor term may be of great relative magnitude, but always has vanishing determinant. The removal of this tensor contribution greatly improves general stability and conserves precision.

All that remains is a multidimensional matrix inversion. This step alone affects algorithm stability. While the inversion cannot be eliminated, the process can be simplified. One observes the following transformation:

$$\vec{v}' = \langle \delta\phi(\vec{z}) \rangle_{\rm exp} \quad \text{and} \quad \boldsymbol{C}' = \langle \vec{\delta f}(\vec{z}) \otimes \tilde{\delta f}(\vec{z}) \rangle_{\rm exp} \quad \longrightarrow \quad \vec{v}' = \boldsymbol{C}'\vec{a}' \qquad \text{(C.5)}$$

$$\delta\vec{f}''(\vec{z}) = \boldsymbol{T}\delta\vec{f}(\vec{z}) \longrightarrow \vec{v}'' = \boldsymbol{T}\vec{v}' = (\boldsymbol{T}\boldsymbol{C}'\tilde{\boldsymbol{T}})(\tilde{\boldsymbol{T}}^{-1}\vec{a}') = \boldsymbol{C}''\vec{a}''$$

for orthogonal transformation $\boldsymbol{T}$ diagonalizing $\boldsymbol{C}''$

Since $\boldsymbol{C}''$ appears diagonalized, equation C.5 gives a prescription for simplifying matrix inversion in the linear regression. Diagonalization does involve a new calculation for orthogonal transformation $\boldsymbol{T}$. The new calculation does not however depend on the inversion of curvature tensor $\boldsymbol{C}$. The salient result is recovery after solution of the result $\vec{a}' = \tilde{\boldsymbol{T}}\vec{a}''$ in the original reference frame, and regardless of whether the transformed solution involved a true inversion or a singular value decomposition. If singular values are found, transformation $\boldsymbol{T}$ does in fact allow for recovery of the null eigenvectors.

One thus finds that recourse to extended precision arithmetic can be avoided in many unstable fit problems. At the cost of modestly increased complexity and a few extra passes over data sets, precision instabilities may be quite elegantly addressed. Singular value decomposition analysis also becomes a natural extension of the general regression solution.

C.2 The cantilever

The cantilever is defined as a beam with rectangular cross-section, hanging off a ledge at a fixed end, with the other end free. The cantilever sags under its own weight and perhaps under an additional end load. Solution of

the cantilever problem derives simply from a most elementary treatment of stress and strain. Complexities of rigorous elastic deformation analysis are not required and actually complicate approximate solution for the deformation field.[BHAT 86] For the following demonstration, consider a cantilever with the following attributes: the beam has length l hanging in the positive direction, horizontally over a ledge at location $x = 0$, with ambient vertical gravitational field g. The cantilever has material density ρ, cross sectional area A, and Young's and shear moduli E and G respectively.

An intuitive approach might involve evaluation of deformation under shear. Cantilever shear modulus deformation under its own weight may be expressed by the following differential equation:

$$\int_x^\infty dt\, w(t) = F = GA\frac{dz}{dx} \quad \longrightarrow \quad GA\frac{d^2z}{dx^2} = -w(x) \qquad \text{(C.6)}$$

$$\text{with boundary conditions } \; z(0) = 0 \text{ and } \frac{dz}{dx}(l) = 0$$

With constant weight density $w = \rho g$, the solution to the shear modulus differential equation may be written:

$$z = \frac{\rho g}{2G}\, x\,(2l - x) \qquad \text{(C.7)}$$

This deflection is constant regardless of applied end load, and should therefore not be observable as a relative deformation between any pair of observations. The effect of end load (mass $= m$) variation should however be observable. By linear superposition of solutions, one may denote this contribution by primed variables, and apply equation C.6 in homogeneous form,

but with new boundary conditions:

$$w(x) = mg\,\delta(l-x) \quad \longrightarrow \quad z'(0) = 0 \text{ and } \frac{dz'}{dx}(l) = +\frac{mg}{GA} \qquad \text{(C.8)}$$

Solution reduces to the elementary form $z' = \frac{mg}{GA}x$. Calculation for the experimental aluminum beam of this analysis shows shear strain to be a negligible effect.

The foregoing evaluation of shear strain models the more complicated effect of Young's modulus strain. Young's modulus strain is longitudinal along the beam. The effect is as though narrow filaments of cantilever stretch quite independent of one another. Under the influence of torque, the beam should bend due to unbalanced stretching of independent filaments. The observable effect of an angular deviation $d\theta/dx$ should follow the elementary Young's modulus form $F = EA\,\delta l/l$ for each independent filament:

$$\tau = \int_A dr = \int_A z \times (E dy dz\, z\frac{d\theta}{dx}) = E\,(\int_A dy dz\, z^2)\,\frac{d\theta}{dx} = EI\,\frac{d\theta}{dx} \qquad \text{(C.9)}$$

The moment I sensibly should be mimimized, as static strains minimize local stresses. The requisite minimum value corresponds to the moment about the cross-section centroid. For a rectangular cross-section of breadth b and depth d this moment is given by $I = \frac{1}{12}bd^3$. The approximation of this derivation, that coordinate x tracks along the locus of cross-section centroids, holds validity for small angle deviations, with corrections of the order $O(\theta^2)$.

To the same order one may utilize the familiar linearization $\frac{dz}{dx} = \tan\theta \sim \theta$ and obtain a Young's modulus equation valid for small angular deviations.

Cantilever Young's modulus deformation under its own weight may be expressed by the following differential equation:[EDWA 85]

$$\int_x^\infty dt\,(t-x)w(t) = \tau = EI\frac{d^2z}{dx^2} \quad \longrightarrow \quad EI\frac{d^4z}{dx^4} = +w(x) \qquad \text{(C.10)}$$

$$\text{with boundary conditions } z(0) = \frac{dz}{dx}(0) = \frac{d^2z}{dx^2}(l) = \frac{d^3z}{dx^3}(l) = 0$$

With constant weight density $w = \rho g$, the solution of the Young's modulus differential equation may be written:

$$z = \frac{\rho g A}{24EI}x^2(x^2 - 4lx + 6l^2) \qquad \text{(C.11)}$$

This deflection is constant regardless of applied end load, and should therefore not be observable as a relative deformation between any pair of observations. The effect of end load variation should however be observable. By linear superposition of solutions, one may denote this contribution by primed variables and apply equation C.10 with a single altered boundary condition:

$$w(x) = mg\,\delta(l-x) \quad \longrightarrow \quad \frac{d^3z}{dx^3}(l) = -\frac{mg}{EI} \qquad \text{(C.12)}$$

Solution for observable Young's modulus deformation may be expressed as the following simple formula:

$$z' = \frac{mg}{6EI}x^2(3l - x) \qquad \text{(C.13)}$$

Amongst observable deformations, Young's modulus deformations dominate over their shear modulus counterparts. Comparision of maximal deflections at the free cantilever end finds the following disparate ratio between

shear and Young's type deflections:

$$R = \frac{z'^{max}_{shear}}{z'^{max}_{Young}} = \frac{1}{4}\frac{E}{G}(\frac{d}{l})^2 \qquad (C.14)$$

The experimental aluminum cantilever of the current research had values $E = 70\,GPa$ (Young's modulus) and $G = 25\,GPa$ (shear modulus) as well as dimensions $l = 10''$ (length), $b = 1''$ (breadth) and $d = 1/4''$ (depth). Calculation of the deflection ratio in equation C.14 gives $R \sim 0.044\%$. Clearly, the shear modulus component amounts to a negligible contribution in the net deflection. The Young's modulus deflection of equation C.13 thus quantifies a model for observable cantilever deformations as examined in this thesis.

The cantilever of this research had an anomaly: a hole had been bored, vertically through through the beam, a short distance into the overhang. The location is apparent in figures 5.14 and 5.15 of section 5.2. Due to complexities involved, values of parameters describing the experimental cantilever were enumerated in a data text *(CNT)* file. Figure C.1 displays precisely the file utilized for all included experimental cantilever deformations, with all spatial units measured in millimetres.

```
9.810000e+03   // gravitational constant
7.000000e+10   // Young's modulus
155.5          // cantilever length
25.4           // cantilever width
6.35           // cantilever depth
15.5           // hole location (from ledge)
3.175          // hole radius
```

Figure C.1: Cantilever model parameters in *CNT* file format

The effect of the hole could be modelled via perturbation expansion of the general integro-differential form of equation C.10 with view only to the observable contribution corresponding to the boundary conditions of equation C.12. Analysis may be motivated by the following expression:

$$mg(l-x) = E(I+\epsilon\delta I)\,(z_0'' + \epsilon z_1'' + \epsilon^2 z_2'') \tag{C.15}$$

Denoting the hole radius by r and its centre location by λ, solution obtains by standard methods, involving a few elementary angular integrations, for perturbative corrections. The identification $z_0''(x) = z'(x)$ already appears in equation C.13. The full solution set, to second order, may be enumerated:

$$\begin{aligned} z_0'' &= \frac{mg}{6EI}\,x^2\,(3l-x) \\ z_1'' &= \pi r^2\,\frac{mg}{b\,EI}\,(l-\lambda)(x-\lambda) \\ z_2'' &= \frac{16}{3}\,r^2\,\frac{mg}{b\,EI}\,(\frac{r}{b})\,[(l-\lambda)(x-\lambda) + \frac{23}{40}\,r^2] \end{aligned} \tag{C.16}$$

Net perturbation solution derives from assertion $\epsilon = 1$ in equation C.15, and yields simply $z'' = z_0'' + z_1'' + z_2''$, with contributions expressed in equation C.16. The perturbations were found to be experimentally significant, and thus cantilever model fits appearing in section 5.3 consistently rely on the complete second order perturbation solution.

Due to the obvious and convenient linear scaling of equations C.15 and C.16, all cantilever model fits utilized applied end load mass m as sensible free parameter. The measured deflection function is so simply described.

Appendix D

Notation

Wherever possible, consistent symbolic notation has been utilized throughout this thesis. Frequently referenced symbols are listed here along with the equations of their definition or extension.

Symbol	Equation	Definition
$(\ldots) * (\ldots)$	A.25	convolution binary operator
$\langle \ldots \rangle$	A.3,A.4	mathematical expectation (unit norm)
$\langle \ldots \rangle_{\text{exp}}$	B.27	experimental expectation (population norm)
$A(z,t)$	A.60	simple plane wave (complex) function
$A_\theta \& A_\theta^*$	A.64	phase offset simple plane wave amplitude
$\vec{A}(\vec{r},t)$	A.73	realistic plane wave (complex) vector potential
$\vec{A}_\theta \& \vec{A}_\theta^*$	A.93	phase offset realistic *EM* plane wave amplitude
$\vec{B}(\vec{r},t)$	A.73	realistic plane wave (complex) magnetic field
$c(\phi;\Phi)$	2.22,A.15	Cauchy density probability function
C_{ij}	3.22	curvature tensor components for *SVD* fit
χ	2.24,3.18,3.30	least-squares & *SVD* minimization functionals
$\chi(x;\nu)$	2.12,A.14	chi-square density probability function

Symbol	Equation	Definition
$\vec{d}(x,y)$	3.4	surface deformation displacement vector field
$\vec{d}_{\text{sing}}(x,y)$	3.9	invisible constructed *singular* displacement
$\vec{d}_{\text{swing}}(x,y)$	3.6	invisible rigid *swing stage* displacement
$\delta\lambda_D$	3.28	systematic decoherent wave dilation effect
$\vec{\Delta}_x$	3.25	phase systematic error generators by cause x
E	C.9	elastic deformation Young's modulus
$\vec{E}(\vec{r},t)$	A.73	realistic plane wave (complex) electric field
$f_k(x,y)$	3.21	fit functions *(SVD)* fit computational view
$f_{nmi}(x,y)$	3.20	fit functions *(SVD)* natural euclidean view
$\mathcal{F}_{x\to k}\{f(x)\}(k)$	A.26	Fourier transform definition
G	C.6	elastic deformation shear modulus
$g(x;\mu,\sigma)$	A.12	Gaussian density probability function
$G_N(\vec{x};\vec{\mu},\vec{\sigma}^2)$	B.17	multivariate Gaussian probability density
$\gamma(x;\alpha,\beta)$	2.1,A.13	gamma density probability function
i	1.7,2.2	single source speckle intensity
$i_\perp \& i_\top$	2.3,A.83	principal axes intensity tensor moments
$i_A \& i_B$	1.4	interfering signal intensities
$I(i_A,i_B;\phi)$	1.4	net interference intensity
$I(\vec{r},t;\hat{n})$	A.76	realistic *EM* wave intensity
$I_\pm(z,t)$	A.60	simple plane wave real/imaginary intensity
$\mathcal{I}(\ldots)$	2.3,A.83	speckle intensity distribution tensor
$\hat{k}$	2.3,A.82	plane-wave propagation ray unit vector
$\vec{k}(x,y)$	3.3	interferometer phase sensitivity vector field
$\kappa(\{i_A\},\{i_B\})$	1.5,A.98	cross-correlation to phase cosine scale
$\mathcal{L}_{x\to s}\{f(x)\}(s)$	A.27	Laplace transform definition

Symbol	Equation	Definition
λ	3.1	coherent laser scatter wavelength
$\mu = \bar{x}$	A.5,A.9	statistical mean (global population)
$\tilde{\mu}$	A.5	statistical median (global,sample)
$\bar{\mu}$	A.8	statistical mode (global population)
$\vec{\mu}$	B.17	multivariate Gaussian mean vector
$\hat{n}$	2.3,A.86	incidence surface normal unit vector
ν	2.12	statistical sample degrees of freedom
$P(i)$	1.7,2.2	speckle intensity probability function
P_{nmi}	3.19	bivariate polynomial displacement coefficients
$\vec{P}_N(x,y)$	3.12,3.19	bivariate polynomial displacement vector field
ϕ	1.4	interference relative phase (theoretical)
ϕ_{dev}	3.29	systematic non-integral deviant phase offset
Φ	3.1	interference relative phase (experimental)
$Q(\rho, \sigma^2_{\rho\rho})$	2.11	sample to linearized $\sigma^2_{\rho\rho}$ variance ratio
$Q_{N-1}(x,y)$	3.12	bivariate polynomial (multiplying) scalar field
$\rho(x,y)$	1.2,1.3,1.9	Pearson's cross-correlation
$\vec{S}(\vec{r},t)$	A.76	Poynting vector for realistic *EM* waves
$\vec{S}_{\pm}(\vec{r},t)$	A.77	Poynting vector for real/imaginary *EM* waves
$\mathcal{S}(\{i_A\},\{i_B\})$	2.5,A.96	speckle intensity distribution tensor
$\sigma^2_{\rho\rho}$	2.9,B.24	cross-correlation sample variance
σ^2_{xx}	A.6,A.10	statistical variance (global,sample)
$\vec{\vec{\sigma}}^2$	B.17	multivariate Gaussian covariance tensor
$\mathcal{T}(\{i_A\},\{i_B\})$	2.5,A.96	speckle intensity distribution tensor
v_i	3.22	data vector components for *SVD* fit
$w(x)$	2.20	data weight for least-squares *SVD* fit